

FILED BY IDS

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-13790

(P2000-13790A)

(43) 公開日 平成12年1月14日 (2000.1.14)

(51) Int.Cl.⁷H04N 7/24
1/41

識別記号

FI

H04N 7/13
1/41

テマコード* (参考)

Z 5C059
B 5C078

BEST AVAILABLE COPY

審査請求 未請求 請求項の数30 O.L. (全 51 頁)

(21) 出願番号 特願平10-173499

(22) 出願日 平成10年6月19日 (1998.6.19)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 名雲 政文

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(72) 発明者 鈴木 輝彦

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(74) 代理人 100082131

弁理士 稲本 義雄

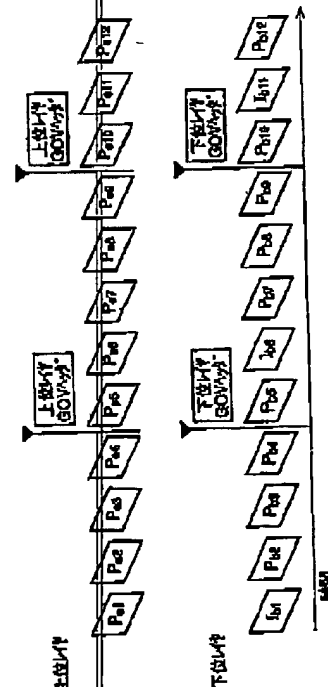
最終頁に続く

(54) 【発明の名称】 画像符号化装置および画像符号化方法、画像復号装置および画像復号方法、並びに提供媒体

(57) 【要約】

【課題】 効率的かつ迅速なランダムアクセスをする。

【解決手段】 画像を構成するオブジェクトのシーケンスが上位レイヤと下位レイヤとに階層化され、下位レイヤのVOP (Video Object Plane) のシーケンスが、GOV (Group Of VOP) に分けて符号化されるとともに、下位レイヤのGOVの最初に表示されるVOPの表示時刻と同時刻またはその直後に表示される上位レイヤのVOPが、GOVの最初に表示されるVOPとなるように、上位レイヤのVOPのシーケンスが、GOVに分けて符号化される。



(2)

特開2000-13790

2

【特許請求の範囲】

【請求項1】 画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であつて、

前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化する階層化手段と、

前記階層化手段が出力する、前記オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化する第1の符号化手段と、

前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、前記階層化手段が出力する、前記オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化する第2の符号化手段とを備えることを特徴とする画像符号化装置。

【請求項2】 前記第1または第2の符号化手段は、前記第1または第2の階層のグループそれぞれに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻を含めることを特徴とする請求項1に記載の画像符号化装置。

【請求項3】 前記第2の符号化手段は、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度と、前記第2の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度とが一致しないとき、前記第2の階層のグループに、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度を、前記秒精度先頭表示時刻として含めることを特徴とする請求項2に記載の画像符号化装置。

【請求項4】 前記第1または第2の符号化手段それぞれは、前記第1または第2の階層のオブジェクトそれぞれに、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することを特徴とする請求項2に記載の画像符号化装置。

【請求項5】 前記第2の符号化手段は、前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報をリセットすることを特徴とする請求項4に記載の画像符号化装置。

【請求項6】 前記階層化手段は、前記画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための前記第1と第2の階層を含む2以上の階層に階層化し、

前記第2の符号化手段は、前記第2の階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することを特徴とする請求項1に記載の画像符号化装置。

【請求項7】 画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化方法であつて、

て、
前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、

前記オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、前記オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することを特徴とする画像符号化方法。

【請求項8】 画像を復号する画像復号装置であつて、前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、

前記オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、前記オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを受信する受信手段と、

前記符号化ビットストリームを復号する復号手段とを備えることを特徴とする画像復号装置。

【請求項9】 前記第1または第2の階層のグループそれぞれには、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻が含まれていることを特徴とする請求項8に記載の画像復号装置。

【請求項10】 前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度と、前記第2の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度とが一致しないとき、

前記第2の階層のグループには、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度が、前記秒精度先頭表示時刻として含まれていることを特徴とする請求項9に記載の画像復号装置。

【請求項11】 前記第1または第2の階層のオブジェクトそれぞれには、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報が付加されていることを特徴とする請求項9に記載の画像復号装置。

【請求項12】 前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報がリセットされていることを特徴とする請求項11に記載の画像復号装置。

【請求項13】 前記画像を構成するオブジェクトのシーケンスが、空間スケーラビリティを実現するための前記第1と第2の階層を含む2以上の階層に階層化されて

(3)

特開 2000-13790

4

おり、

前記第2の階層のオブジェクトのシーケンスが、その表示順序と同一の順序で符号化されていることを特徴とする請求項8に記載の画像復号装置。

【請求項14】 画像を復号する画像復号方法であって、

前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、前記オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、前記オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを受信し、前記符号化ビットストリームを復号することを特徴とする画像復号方法。

【請求項15】 画像を符号化して得られる符号化ビットストリームを提供する提供媒体であって、前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、前記オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、前記オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる前記符号化ビットストリームを提供することを特徴とする提供媒体。

【請求項16】 前記第1または第2の階層のグループそれぞれには、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻が含まれていることを特徴とする請求項15に記載の提供媒体。

【請求項17】 前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度と、前記第2の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度とが一致しないとき、前記第2の階層のグループには、前記第1の階層のグループの最初に表示されるオブジェクトの表示時刻の秒精度が、前記秒精度先頭表示時刻として含まれていることを特徴とする請求項16に記載の提供媒体。

【請求項18】 前記第1または第2の階層のオブジェクトそれぞれには、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報が付加されていることを特徴とする請求項16に記載の提供媒体。

【請求項19】 前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第

2の階層のオブジェクトについての前記秒精度相対時刻情報がリセットされていることを特徴とする請求項18に記載の提供媒体。

【請求項20】 前記画像を構成するオブジェクトのシーケンスが、空間スケーラビリティを実現するための前記第1と第2の階層を含む2以上の階層に階層化されており、

前記第2の階層のオブジェクトのシーケンスが、その表示順序と同一の順序で符号化されていることを特徴とする請求項15に記載の提供媒体。

【請求項21】 画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、

前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化する階層化手段と、

前記階層化手段が出力する、前記オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、前記第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含める符号化手段と、

前記第1または第2の階層のオブジェクトそれぞれに、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加する付加手段と、

前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報をリセットするリセット手段とを備えることを特徴とする画像符号化装置。

【請求項22】 画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化方法であって、

前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、

前記オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、前記第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、

前記第1または第2の階層のオブジェクトそれぞれに、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加する画像符号化方法において、

前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報をリセットすることを特徴とする画像符号化方法。

【請求項23】 画像を復号する画像復号装置であって、

50

(4)

特開2000-13790

5

前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、

前記オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、前記第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、

前記第1または第2の階層のオブジェクトそれぞれに、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報がリセットされているものを受信する受信手段と、前記符号化ビットストリームを復号する復号手段とを備えることを特徴とする画像復号装置。

【請求項24】 画像を復号する画像復号方法であって、

前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、

前記オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、前記第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、

前記第1または第2の階層のオブジェクトそれぞれに、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、前記第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報がリセットされているものを受信し、前記符号化ビットストリームを復号することを特徴とする画像復号方法。

【請求項25】 画像を符号化して得られる符号化ビットストリームを提供する提供媒体であって、前記画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、

前記オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、前記第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、

前記第1または第2の階層のオブジェクトそれぞれに、前記秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、前記第2の階層の、表示順で隣接するオブジェクト

6

どうしの表示時刻の差に基づいて、前記第2の階層のオブジェクトについての前記秒精度相対時刻情報がリセットされているものを提供することを特徴とする提供媒体。

【請求項26】 画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、

前記画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化する階層化手段と、前記階層化手段が出力する、前記下位階層のオブジェクトのシーケンスを符号化する第1の符号化手段と、前記階層化手段が出力する、前記上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化する第2の符号化手段とを備えることを特徴とする画像符号化装置。

【請求項27】 画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化方法であって、

前記画像を受信し、前記画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、前記下位階層のオブジェクトのシーケンスを符号化するとともに、前記上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することを特徴とする画像符号化方法。

【請求項28】 画像を復号する画像復号装置であって、

前記画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、前記下位階層のオブジェクトのシーケンスを符号化するとともに、前記上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを受信する受信手段と、

前記符号化ビットストリームを復号する復号手段とを備えることを特徴とする画像復号装置。

【請求項29】 画像を復号する画像復号方法であって、

前記画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、

前記下位階層のオブジェクトのシーケンスを符号化するとともに、前記上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを受信し、

前記符号化ビットストリームを復号することを特徴とする画像復号方法。

(5)

特開2000-13790

7

【請求項30】 画像を符号化して得られる符号化ビットストリームを提供する提供媒体であって、前記画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、前記下位階層のオブジェクトのシーケンスを符号化するとともに、前記上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを提供することを特徴とする提供媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、画像符号化装置および画像符号化方法、画像復号装置および画像復号方法、並びに提供媒体に関する。特に、例えば、動画像データを、光磁気ディスクや磁気テープなどの記録媒体に記録し、これを再生してディスプレイなどに表示したり、テレビ会議システム、テレビ電話システム、放送用機器、マルチメディアデータベース検索システムなどのように、動画像データを伝送路を介して送信側から受信側に伝送し、受信側において、受信された動画像データを表示する場合や、編集して記録する場合などに用いて好適な画像符号化装置および画像符号化方法、画像復号装置および画像復号方法、並びに提供媒体に関する。

【0002】

【従来の技術】例えば、テレビ会議システム、テレビ電話システムなどのように、動画像データを遠隔地に伝送するシステムにおいては、伝送路を効率良く利用するため、画像データを、そのライン相関やフレーム間相関を利用して圧縮符号化するようになされている。

【0003】動画像の高効率符号化方式として代表的なものとしては、MPEG (Moving Picture Experts Group) (著信用動画像符号化) 方式がある。これはISO-IEC/JTC1/SC2/WG11において議論され、標準案として提案されたものであり、動き補償予測符号化とDCT (Discrete Cosine Transform) 符号化を組み合わせたハイブリッド方式が採用されている。

【0004】MPEGでは、様々なアプリケーションや機能に対応するために、いくつかのプロファイルおよびレベルが定義されている。最も基本となるのが、メインプロファイルメインレベル (MP@ML (Main Profile at Main Level)) である。

【0005】図53は、MPEG方式におけるMP@MLのエンコーダの一例の構成を示している。

【0006】符号化すべき画像データは、フレームメモリ31に入力され、一時記憶される。そして、動きベクトル検出器32は、フレームメモリ31に記憶された画像データを、例えば、16画素×16画素などで構成されるマクロブロック単位で読み出し、その動きベクトルを検出する。

8

【0007】ここで、動きベクトル検出器32においては、各フレームの画像データを、Iピクチャ(フレーム内符号化)、Pピクチャ(前方予測符号化)、またはBピクチャ(両方向予測符号化)のうちのいずれかとして処理する。なお、シーケンシャルに入力される各フレームの画像を、I、P、Bピクチャのいずれのピクチャとして処理するかは、例えば、予め定められている(例えば、I、B、P、B、P、・・・B、Pとして処理される)。

10 【0008】即ち、動きベクトル検出器32は、フレームメモリ31に記憶された画像データの中の、予め定められた所定の参照フレームを参照し、その参照フレームと、現在符号化の対象となっているフレームの16画素×16ラインの小ブロック(マクロブロック)とをパターンマッチング(ブロックマッチング)することにより、そのマクロブロックの動きベクトルを検出する。

20 【0009】ここで、MPEGにおいては、画像の予測モードには、イントラ符号化(フレーム内符号化)、前方予測符号化、後方予測符号化、両方向予測符号化の4種類があり、Iピクチャはイントラ符号化され、Pピクチャはイントラ符号化または前方予測符号化のいずれかで符号化され、Bピクチャはイントラ符号化、前方予測符号化、後方予測符号化、または両方法予測符号化のいずれかで符号化される。

30 【0010】即ち、動きベクトル検出器32は、Iピクチャについては、予測モードとしてイントラ符号化モードを設定する。この場合、動きベクトル検出器32は、動きベクトルの検出は行わず、予測モード(イントラ予測モード)を、VLC(可変長符号化)器36および動き補償器42に出力する。

40 【0011】また、動きベクトル検出器32は、Pピクチャについては、前方予測を行い、その動きベクトルを検出する。さらに、動きベクトル検出器32は、前方予測を行うことにより生じる予測誤差と、符号化対象のマクロブロック(Pピクチャのマクロブロック)の、例えば分散とを比較する。その比較の結果、マクロブロックの分散の方が予測誤差より小さい場合、動きベクトル検出器32は、予測モードとしてイントラ符号化モードを設定し、VLC器36および動き補償器42に出力する。また、動きベクトル検出器32は、前方予測を行うことにより生じる予測誤差の方が小さければ、予測モードとして前方予測符号化モードを設定し、検出した動きベクトルとともに、VLC器36および動き補償器42に出力する。

50 【0012】さらに、動きベクトル検出器32は、Bピクチャについては、前方予測、後方予測、および両方向予測を行い、それぞれの動きベクトルを検出する。そして、動きベクトル検出器32は、前方予測、後方予測、および両方向予測についての予測誤差の中の最小のものを(以下、適宜、最小予測誤差という)を検出し、その最

(6)

特開2000-13790

9

10

小予測誤差と、符号化対象のマクロブロック（Bピクチャのマクロブロック）の、例えば分散とを比較する。その比較の結果、マクロブロックの分散の方が最小予測誤差より小さい場合、動きベクトル検出器32は、予測モードとしてイントラ符号化モードを設定し、VLC器36および動き補償器42に出力する。また、動きベクトル検出器32は、最小予測誤差の方が小さければ、予測モードとして、その最小予測誤差が得られた予測モードを設定し、対応する動きベクトルとともに、VLC器36および動き補償器42に出力する。

【0013】動き補償器42は、動きベクトル検出器32から予測モードと動きベクトルの両方を受信すると、その予測モードおよび動きベクトルにしたがって、フレームメモリ41に記憶されている、符号化され、かつ既に局所復号された画像データを読み出し、この読み出された画像データを、予測画像データとして、演算器33および40に供給する。

【0014】演算器33は、動きベクトル検出器32がフレームメモリ31から読み出した画像データと同一のマクロブロックをフレームメモリ31から読み出し、そのマクロブロックと、動き補償器42からの予測画像との差分を演算する。この差分値は、DCT器34に供給される。

【0015】一方、動き補償器42は、動きベクトル検出器32から予測モードのみを受信した場合、即ち、予測モードがイントラ符号化モードである場合には、予測画像を出力しない。この場合、演算器33（演算器40も同様）は、特に処理を行わず、フレームメモリ31から読み出したマクロブロックを、そのままDCT器34に出力する。

【0016】DCT器34では、演算器33の出力データに対して、DCT処理が施され、その結果得られるDCT係数が、量子化器35に供給される。量子化器35では、バッファ37のデータ蓄積量（バッファ37に記憶されているデータの量）（バッファフィードバック）に対応して量子化ステップ（量子化スケール）が設定され、その量子化ステップで、DCT器34からのDCT係数が量子化される。この量子化されたDCT係数（以下、適宜、量子化係数という）は、設定された量子化ステップとともに、VLC器36に供給される。

【0017】VLC器36では、量子化器35より供給される量子化係数が、例えばハフマン符号などの可変長符号に変換され、バッファ37に出力される。さらに、VLC器36は、量子化器35からの量子化ステップ、動きベクトル検出器32からの予測モード（イントラ符号化（画像内予測符号化）、前方予測符号化、後方予測符号化、または両方向予測符号化のうちのいずれが設定されたかを示すモード）および動きベクトルも可変長符号化し、その結果得られる符号化ビットストリームを、バッファ37に出力する。

【0018】バッファ37は、VLC器36からの符号化ビットストリームを一時蓄積することにより、そのデータを平滑化し、例えば、伝送路に出力し、または記録媒体に記録する。

【0019】また、バッファ37は、そのデータ蓄積量を量子化器35に出力しており、量子化器35は、このバッファ37からのデータ蓄積量にしたがって量子化ステップを設定する。即ち、量子化器35は、バッファ37がオーバーフローしそうなどとき、量子化ステップを大きくし、これにより、量子化係数のデータ量を低下させる。また、量子化器35は、バッファ37がアンダーフローしそうなどとき、量子化ステップを小さくし、これにより、量子化係数のデータ量を増大させる。このようにして、バッファ37のオーバーフローとアンダーフローを防止するようになっている。

【0020】量子化器35が出力する量子化係数と量子化ステップは、VLC器36だけでなく、逆量子化器38にも供給されるようになされている。逆量子化器38では、量子化器35からの量子化係数が、同じく量子化器35からの量子化ステップにしたがって逆量子化され、これによりDCT係数に変換される。このDCT係数は、IDCT器（逆DCT器）39に供給される。IDCT器39では、DCT係数が逆DCT処理され、その処理の結果得られるデータが、演算器40に供給される。

【0021】演算器40には、IDCT器39の出力データの他、上述したように、動き補償器42から、演算器33に供給されている予測画像と同一のデータが供給されている。演算器40は、IDCT器39の出力データ（予測残差（差分データ））と、動き補償器42からの予測画像データとを加算することで、元の画像データを局所復号し、この局所復号された画像データ（局所復号画像データ）が出力される（但し、予測モードがイントラ符号化である場合には、IDCT器39の出力データは、演算器40をスルーして、そのまま、局所復号画像データとして、フレームメモリ41に供給される）。なお、この復号画像データは、受信側において得られる復号画像データと同一のものである。

【0022】演算器40において得られた復号画像データ（局所復号画像データ）は、フレームメモリ41に供給されて記憶され、その後、インター符号化（前方予測符号化、後方予測符号化、量方向予測符号化）される画像に対する参照画像データ（参照フレーム）として用いられる。

【0023】次に、図54は、図53のエンコーダから出力される符号化ビットストリームを復号する、MP@MLにおけるMP@MLのデコーダの一例の構成を示している。

【0024】伝送路を介して伝送されてきた符号化ビットストリームが図示せぬ受信装置で受信され、または記

(7)

特開2000-13790

11

録媒体に記録された符号化ビットストリームが図示せぬ再生装置で再生され、バッファ101に供給されて記憶される。

【0025】IVLC器(逆VLC器(可変長復号器))102は、バッファ101に記憶された符号化ビットストリームを読み出し、可変長復号することにより、その符号化ビットストリームを、マクロブロック単位で、動きベクトル、予測モード、量子化ステップ、および量子化係数に分離する。これらのデータのうち、動きベクトルおよび予測モードは動き補償器107に供給され、量子化ステップおよびマクロブロックの量子化係数は逆量子化器103に供給される。

【0026】逆量子化器103は、IVLC器102より供給されたマクロブロックの量子化係数を、同じくIVLC器102より供給された量子化ステップにしたがって逆量子化し、その結果得られるDCT係数を、IDCT器104に出力する。IDCT器104は、逆量子化器103からのマクロブロックのDCT係数を逆DCTし、演算器105に供給する。

【0027】演算器105には、IDCT器104の出力データの他、動き補償器107の出力データも供給されている。即ち、動き補償器107は、フレームメモリ106に記憶されている、既に復号された画像データを、図53の動き補償器42における場合と同様に、IVLC器102からの動きベクトルおよび予測モードにしたがって読み出し、予測画像データとして、演算器105に供給する。演算器105は、IDCT器104の出力データ(予測残差(差分値))と、動き補償器107からの予測画像データとを加算することで、元の画像データを復号する。この復号画像データは、フレームメモリ106に供給されて記憶される。なお、IDCT器104の出力データが、イントラ符号化されたものである場合には、その出力データは、演算器105をスルーして、復号画像データとして、そのままフレームメモリ106に供給されて記憶される。

【0028】フレームメモリ106に記憶された復号画像データは、その後に復号される画像データの参照画像データとして用いられる。さらに、復号画像データは、出力再生画像として、例えば、図示せぬディスプレイなどに供給されて表示される。

【0029】なお、MPEG1および2では、Bピクチャは、参照画像データとして用いられないため、エンコードまたはデコードのそれぞれにおいて、フレームメモリ41(図53)または106(図54)には記憶されない。

【0030】

【発明が解決しようとする課題】以上の図53、図54に示したエンコード、デコードは、MPEG1/2の規格に準拠したものであるが、現在、画像を構成する物体などのオブジェクトのシーケンスであるVO (Video Ob

12

ject) 単位で符号化を行う方式につき、ISO-IEC/JTC1/SC29/WG11において、MPEG (Moving Picture Experts Group) 4として標準化作業が進められている。

【0031】ところで、MPEG4については、主として、通信の分野で利用されるものとして、標準化作業が進められていたため、MPEG1/2において規定されているGOP (Group Of Picture) は、当初、MPEG4では規定されておらず、従って、MPEG4が蓄積メディアに利用された場合には、効率的なランダムアクセスが困難になることが予想される。

【0032】このため、本件出願人は、効率的なランダムアクセスを可能とするために、MPEG1/2で規定されているGOPに相当するGOV (Group Of VOP)の導入を、特願平10-80758号において先に提案しており、また、MPEG4において、このGOVが導入された。

【0033】一方、MPEG4では、画像データを2以上の階層に階層化し、各階層の画像を利用した、柔軟なスケラブル符号化/復号が可能となっている。

【0034】ところで、MPEG4では、現在、各階層の画像データのGOVどうしの対応関係を規定しておらず、このため、各階層ごとに、独立に、GOVを挿入することが可能である。しかしながら、各階層の画像データは、独立したものではないから、各階層ごとに、独立に、GOVを挿入した場合には、効率的なランダムアクセスが困難となる場合が生じることが予想される。

【0035】本発明は、このような状況に鑑みてなされたものであり、効率的なランダムアクセスをすることができるようになるものである。

【0036】

【課題を解決するための手段】請求項1に記載の画像符号化装置は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化する階層化手段と、階層化手段が出力する、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化する第1の符号化手段と、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、階層化手段が出力する、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化する第2の符号化手段とを備えることを特徴とする。

【0037】請求項7に記載の画像符号化方法は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェク

(8)

特開 2000-13790

13

14

トとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することを特徴とする。

【0038】請求項8に記載の画像復号装置は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを復号する復号手段を備えることを特徴とする。

【0039】請求項14に記載の画像復号方法は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを復号することを特徴とする。

【0040】請求項15に記載の提供媒体は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを提供することを特徴とする。

【0041】請求項21に記載の画像符号化装置は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化する階層化手段と、階層化手段が出力する、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含める符号化手段と、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加する付加手段と、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報をリセットするリセット手段とを備えるこ

とを特徴とする。

【0042】請求項22に記載の画像符号化方法は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加する画像符号化方法において、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報をリセットすることを特徴とする。

【0043】請求項23に記載の画像復号装置は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを復号する復号手段を備えることを特徴とする。

【0044】請求項24に記載の画像復号方法は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを復号することを特徴とする。

【0045】請求項25に記載の提供媒体は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示され

るオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを提供することを特徴とする。

【0046】請求項26に記載の画像符号化装置は、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化する階層化手段と、階層化手段が出力する、下位階層のオブジェクトのシーケンスを符号化する第1の符号化手段と、階層化手段が出力する、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化する第2の符号化手段とを備えることを特徴とする。

【0047】請求項27に記載の画像符号化方法は、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することを特徴とする。

【0048】請求項28に記載の画像復号装置は、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを復号する復号手段を備えることを特徴とする。

【0049】請求項29に記載の画像復号方法は、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを復号することを特徴とする。

【0050】請求項30に記載の提供媒体は、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを提供することを特徴とする。

【0051】請求項1に記載の画像符号化装置において

は、階層化手段は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、第1の符号化手段は、階層化手段が出力する、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するようになされている。第2の符号化手段は、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、階層化手段が出力する、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化するようになされている。

【0052】請求項7に記載の画像符号化方法においては、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化するようになされている。

【0053】請求項8に記載の画像復号装置においては、復号手段が、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを復号するようになされている。

【0054】請求項14に記載の画像復号方法においては、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを復号するようになされている。

【0055】請求項15に記載の提供媒体においては、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブ

(10)

特開 2000-13790

17

エクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームを提供するようになされている。

【0056】請求項21に記載の画像符号化装置においては、階層化手段は、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、符号化手段は、階層化手段が出力する、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含めるようになされている。付加手段は、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加し、リセット手段は、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報をリセットするようになされている。

【0057】請求項22に記載の画像符号化方法においては、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加するようになされている。この場合に、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報をリセットするようになされている。

【0058】請求項23に記載の画像復号装置においては、復号手段が、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを復号するようになされている。

【0059】請求項24に記載の画像復号方法においては、画像を構成するオブジェクトのシーケンスを2以上

18

の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを復号するようになされている。

【0060】請求項25に記載の提供媒体においては、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを提供するようになされている。

【0061】請求項26に記載の画像符号化装置においては、階層化手段は、画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、第1の符号化手段は、階層化手段が出力する、下位階層のオブジェクトのシーケンスを符号化するようになされている。第2の符号化手段は、階層化手段が出力する、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化するようになされている。

【0062】請求項27に記載の画像符号化方法においては、画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化するようになされている。

【0063】請求項28に記載の画像復号装置においては、復号手段が、画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序

(11)

特開 2000-13790

19

と同一の順序で符号化することにより得られる符号化ビットストリームを復号するようになされている。

【0064】請求項29に記載の画像復号方法においては、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを表現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを復号するようになされている。

【0065】請求項30に記載の提供媒体においては、画像を構成するオブジェクトのシーケンスを、空間スケラビリティを表現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを提供するようになされている。

【0066】

【発明の実施の形態】以下に、本発明の実施の形態を説明するが、その前に、特許請求の範囲に記載の発明の各手段と以下の実施の形態との対応関係を明らかにするために、各手段の後の括弧内に、対応する実施の形態（但し、一例）を付加して、本発明の特徴を記述すると、次のようになる。

【0067】即ち、請求項1に記載の画像符号化装置は、画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化する階層化手段（例えば、図3に示す画像階層化部21など）と、階層化手段が出力する、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化する第1の符号化手段（例えば、図3に示す下位レイヤ符号化部25など）と、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、階層化手段が出力する、オブジェクトのシーケンスの第2の階層を、複数のグループに分けて符号化する第2の符号化手段（例えば、図3に示す上位レイヤ符号化部23など）とを備えることを特徴とする。

【0068】請求項8に記載の画像復号装置は、画像を復号する画像復号装置であって、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1の階層を、複数のグループに分けて符号化するとともに、第1の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第2の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第2の階層を、複数のグル

20

ープに分けて符号化することにより得られる符号化ビットストリームを受信する受信手段（例えば、図14に示す逆多重化部91など）と、符号化ビットストリームを復号する復号手段（例えば、図14に示す上位レイヤ復号部93および下位レイヤ復号部95など）とを備えることを特徴とする。

【0069】請求項21に記載の画像符号化装置は、画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化する階層化手段（例えば、図3に示す画像階層化部21など）と、階層化手段が出力する、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含める符号化手段（例えば、図3に示す上位レイヤ符号化部23および下位レイヤ符号化部25など）と、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加する付加手段（例えば、図3に示す上位レイヤ符号化部23および下位レイヤ符号化部25など）と、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報をリセットするリセット手段（例えば、図48に示すプログラムの処理ステップS23など）とを備えることを特徴とする。

【0070】請求項23に記載の画像復号装置は、画像を復号する画像復号装置であって、画像を構成するオブジェクトのシーケンスを2以上の階層に階層化し、オブジェクトのシーケンスの第1または第2の階層を、1以上のグループに分けてそれぞれ符号化し、第1または第2の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第1または第2の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものを受信する受信手段（例えば、図14に示す逆多重化部91など）と、符号化ビットストリームを復号する復号手段（例えば、図14に示す上位レイヤ復号部93および下位レイヤ復号部95など）とを備えることを特徴とする。

【0071】請求項26に記載の画像符号化装置は、画像を符号化し、その結果得られる符号化ビットストリームを出力する画像符号化装置であって、画像を構成する

(12)

特開 2000-13790

21

オブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化する階層化手段（例えば、図3に示す画像階層化部21など）と、階層化手段が出力する、下位階層のオブジェクトのシーケンスを符号化する第1の符号化手段（例えば、図3に示す下位レイヤ符号化部25など）と、階層化手段が出力する、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化する第2の符号化手段（例えば、図3に示す上位レイヤ符号化部23など）とを備えることを特徴とする。

【0072】請求項28に記載の画像復号装置は、画像を復号する画像復号装置であって、画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームを受信する受信手段（例えば、図14に示す逆多重化部91など）と、符号化ビットストリームを復号する復号手段（例えば、図14に示す上位レイヤ復号部93および下位レイヤ復号部95など）とを備えることを特徴とする。

【0073】なお、勿論この記載は、各手段を上記したものに限定することを意味するものではない。

【0074】図1は、本発明を適用したエンコーダの一実施の形態の構成例を示している。なお、このエンコーダは、基本的には、MPEG4の規格に即した処理を行うようになされている。

【0075】符号化すべき画像（動画像）データは、VO（Video Object）構成部1に入力され、VO構成部1では、そこに入力される画像を構成するオブジェクトごとに、そのシーケンスであるVOが構成され、VOP構成部2、乃至2_nに出力される。即ち、VO構成部1においてN個のVO#1乃至VO#Nが構成された場合、そのN個のVO#1乃至VO#Nは、VOP構成部2、乃至2_nにそれぞれ出力される。

【0076】具体的には、例えば、符号化すべき画像データが、独立した背景F1のシーケンスと前景F2のシーケンスとから構成される場合、VO構成部1は、例えば、前景F2のシーケンスを、VO#1として、VOP構成部2₁に出力するとともに、背景F1のシーケンスを、VO#2として、VOP構成部2₂に出力する。

【0077】なお、VO構成部1は、符号化すべき画像データが、例えば、背景F1と前景F2とを、既に合成したものである場合、所定のアルゴリズムにしたがって、画像を領域分割することにより、背景F1と前景F2とを取り出し、それぞれのシーケンスとしてのVOを、対応するVOP構成部2_i（但し、 $i=1, 2, \dots, N$ ）に出力する。

【0078】VOP構成部2_iは、VO構成部1の出力

22

から、VOP（VO Plane）を構成する。即ち、例えば、各フレームから物体を抽出し、その物体を囲む、例えば、最小の長方形（以下、適宜、最小長方形という）をVOPとする。なお、このとき、VOP構成部2_iは、その横および縦の画素数が、例えば、16の倍数となるようにVOPを構成する。VO構成部2_iは、VOPを構成すると、そのVOPを、VOP符号化部3_iに出力する。

【0079】さらに、VOP構成部2_iは、VOPの大きさ（例えば、横および縦の長さ）を表すサイズデータ（VOP size）と、フレームにおける、そのVOPの位置（例えば、フレームの最も左上を原点とするときの座標）を表すオフセットデータ（VOP offset）とを検出し、これらのデータも、VOP符号化部3_iに供給する。

【0080】VOP符号化部3_iは、VOP構成部2_iの出力を、例えば、MPEGや、H. 263などの規格に準拠した方式で符号化し、その結果得られるビットストリームを、多重化部4に出力する。多重化部4は、VOP符号化部3_i乃至3_nからのビットストリームを多重化し、その結果得られる多重化データを、例えば、地上波や、衛星回線、CATV網その他の伝送路5を介して伝送し、または、例えば、磁気ディスク、光磁気ディスク、光ディスク、磁気テープその他の記録媒体6に記録する。

【0081】ここで、VOおよびVOPについて説明する。

【0082】VOは、ある合成画像のシーケンスが存在する場合、その合成画像を構成する各オブジェクト（物体）のシーケンスであり、VOPは、ある時刻におけるVOを意味する。即ち、例えば、いま、画像F1およびF2を合成して構成される合成画像F3がある場合、画像F1またはF2が時系列に並んだものが、それぞれVOであり、ある時刻における画像F1またはF2が、それぞれVOPである。従って、VOは、異なる時刻の、同一オブジェクトのVOPの集合といえることができる。

【0083】なお、例えば、画像F1を背景とするとともに、画像F2を前景とすると、合成画像F3は、画像F2を抜くためのキー信号を用いて、画像F1およびF2を合成することによって得られるが、この場合における画像F2のVOPには、その画像F2を構成する画像データ（輝度信号および色差信号）の他、適宜、そのキー信号も含まれるものとする。

【0084】画像フレーム（画枠）のシーケンスは、その大きさおよび位置のいずれも変化しないが、VOは、大きさや位置が変化する場合がある。即ち、同一のVOを構成するVOPであっても、時刻によって、その大きさや位置が異なる場合がある。

【0085】具体的には、図2は、背景である画像F1

(13)

特開 2000-13790

23

と、前景である画像F2とからなる合成画像を示している。

【0086】画像F1は、例えば、ある自然の風景を撮影したものであり、その画像全体のシーケンスが1つのVO（VO#0とする）とされている。また、画像F2は、例えば、人が歩いている様子を撮影したものであり、その人を囲む最小の長方形のシーケンスが1つのVO（VO#1とする）とされている。

【0087】この場合、VO#0は風景の画像であるから、基本的に、通常の画像のフレームと同様に、その位置および大きさの両方とも変化しない。これに対して、VO#1は人の画像であるから、人物が左右に移動したり、また、図面において手前側または奥側に移動することにより、その大きさや位置が変化する。従って、図2は、同一時刻におけるVO#0およびVO#1を表しているが、VOの位置や大きさは、時間の経過にともなうて変化することがある。

【0088】そこで、図1のVOP符号化部3は、その出力するビットストリームに、VOPを符号化したデータの他、所定の絶対座標系におけるVOPの位置（座標）および大きさに関する情報も含めるようになされている。なお、図2においては、VO#0を構成する、ある時刻のVOP（画像F1）の位置を示すベクトルをOST0と、その時刻と同一時刻における、VO#1のVOP（画像F2）の位置を表すベクトルをOST1と、それぞれ表してある。

【0089】次に、図3は、スケーラビリティを実現する、図1のVOP符号化部3の構成例を示している。即ち、MPEG4では、異なる画像サイズやフレームレートに対応するスケーラビリティを実現するスケーラブル符号化方式が導入されており、図3に示したVOP符号化部3では、そのようなスケーラビリティを実現することができるようになされている。

【0090】VOP構成部2からのVOP（画像データ）、並びにそのサイズデータ（VOPsize）、およびオフセットデータ（VOP offset）は、いずれも画像階層化部21に供給される。

【0091】画像階層化部21は、VOPから、1以上の階層の画像データを生成する（VOPの1以上の階層化を行う）。即ち、例えば、空間スケーラビリティの符号化を行う場合においては、画像階層化部21は、そこに入力される画像データを、そのまま上位レイヤ（上位階層）の画像データとして出力するとともに、それらの画像データを構成する画素数を間引くことなどにより縮小（解像度を低下させ）、これを下位レイヤ（下位階層）の画像データとして出力する。

【0092】なお、入力されたVOPを下位レイヤのデータとするとともに、そのVOPの解像度を、何らかの手法で高くし（画素数を多くし）、これを、上位レイヤのデータとすることなども可能である。

24

【0093】また、階層数は、1とすることが可能であるが、この場合、スケーラビリティは実現されない。なお、この場合、VOP符号化部3は、例えば、下位レイヤ符号化部25だけで構成されることになる。

【0094】さらに、階層数は、3以上とすることも可能であるが、ここでは、簡単のために、2階層の場合について説明を行う。

【0095】画像階層化部21は、例えば、時間スケーラビリティ（テンポラルスケーラビリティ）の符号化を行う場合、時刻に応じて、画像データを、下位レイヤまたは上位レイヤのデータとして、例えば、交互に出力する。即ち、例えば、画像階層化部21は、そこに、あるVOを構成するVOPが、VOP0、VOP1、VOP2、VOP3、・・・の順で入力されたとした場合、VOP0、VOP2、VOP4、VOP6、・・・を、下位レイヤのデータとして、また、VOP1、VOP3、VOP5、VOP7、・・・を、上位レイヤデータとして出力する。なお、時間スケーラビリティの場合は、このようにVOPが間引かれたものが、下位レイヤおよび上位レイヤのデータとされるだけで、画像データの拡大または縮小（解像度の変換）は行われぬ（但し、行うようにすることも可能である）。

【0096】また、画像階層化部21は、例えば、SNR（Signal to Noise Ratio）スケーラビリティの符号化を行う場合、入力された画像データを、そのまま上位レイヤまたは下位レイヤのデータそれぞれとして出力する。即ち、この場合、下位レイヤ並びに上位レイヤの画像データは、同一のデータとなる。

【0097】ここで、VOPごとに符号化を行う場合の空間スケーラビリティについては、例えば、次のような3種類が考えられる。

【0098】即ち、例えば、いま、VOPとして、図2に示したような画像F1およびF2でなる合成画像が入力されたとなると、第1の空間スケーラビリティは、図4に示すように、入力されたVOP全体（図4（A））を上位レイヤ（Enhancement Layer）とするとともに、そのVOP全体を縮小したもの（図4（B））を下位レイヤ（Base Layer）とするものである。

【0099】また、第2の空間スケーラビリティは、図5に示すように、入力されたVOPを構成する一部の物体（図5（A）（ここでは、画像F2に相当する部分））を抜き出して（なお、このような抜き出しは、例えば、VOP構成部2における場合と同様にして行われ、従って、これにより抜き出された物体も、1つのVOPと考えることができる）、上位レイヤとするとともに、そのVOP全体を縮小したもの（図5（B））を下位レイヤとするものである。

【0100】さらに、第3の空間スケーラビリティは、図6および図7に示すように、入力されたVOPを構成する物体（VOP）を抜き出して、その物体ごとに、上

(14)

特開 2000-13790

25

位レイヤおよび下位レイヤを生成するものである。なお、図6は、図2のVOPを構成する背景（画像F1）から上位レイヤおよび下位レイヤを生成した場合を示しており、また、図7は、図2のVOPを構成する前景（画像F2）から上位レイヤおよび下位レイヤを生成した場合を示している。

【0101】以上のようなスケラビリティのうちのいずれを用いるかは予め決められており、画像階層化部21は、その予め決められたスケラビリティによる符号化を行うことができるように、VOPの階層化を行う。

【0102】さらに、画像階層化部21は、そこに入力されるVOPのサイズデータおよびオフセットデータ（それぞれを、以下、適宜、初期サイズデータ、初期オフセットデータという）から、生成した下位レイヤおよび上位レイヤのVOPの所定の絶対座標系における位置を表すオフセットデータと、その大きさを示すサイズデータとを計算（決定）する。

【0103】ここで、下位レイヤ並びに上位レイヤのVOPのオフセットデータ（位置情報）およびサイズデータの決定方法について、例えば、上述の第2のスケラビリティ（図5）を行う場合を例に説明する。

【0104】この場合、下位レイヤのオフセットデータFPOS_Bは、例えば、図8（A）に示すように、下位レイヤの画像データを、その解像度および上位レイヤの解像度の違いに基づいて拡大（アップサンプリング）したときに、即ち、下位レイヤの画像を、上位レイヤの画像の大きさと一致するような拡大率（上位レイヤの画像を縮小して下位レイヤの画像を生成したときの、その縮小率の逆数）（以下、適宜、倍率FRという）で拡大したときに、その拡大画像の絶対座標系におけるオフセットデータが、初期オフセットデータと一致するように決定される。また、下位レイヤのサイズデータFSZ_Bも同様に、下位レイヤの画像を倍率FRで拡大したときに得られる拡大画像のサイズデータが初期サイズデータと一致するように決定される。即ち、オフセットデータFPOS_BまたはサイズデータFSZ_Bは、それぞれのFR倍か、初期オフセットデータまたは初期サイズデータと一致するように決定される。

【0105】一方、上位レイヤのオフセットデータFPOS_Eは、例えば、図8（B）に示すように、入力されたVOPから抜き出した物体を囲む最小長方形（VOP）の、例えば、左上の頂点の座標が、初期オフセットデータに基づいて求められ、この値に決定される。また、上位レイヤのサイズデータFPOS_Eは、入力されたVOPから抜き出した物体を囲む最小長方形の、例えば横および縦の長さに決定される。

【0106】従って、この場合、下位レイヤのオフセットデータFPOS_BおよびサイズデータFPOS_Bを、倍率FRにしたがって変換し（変換後のオフセットデータFPOS_BまたはサイズデータFPOS_B

26

を、それぞれ、変換オフセットデータFPOS_Bまたは変換サイズデータFPOS_Bという）、絶対座標系において、変換オフセットデータFPOS_Bに対応する位置に、変換サイズデータFSZ_Bに対応する大きさの画枠を考え、そこに、下位レイヤの画像データをFR倍だけした拡大画像を配置するとともに（図8

（A））、その絶対座標系において、上位レイヤのオフセットデータFPOS_EおよびサイズデータFPOS_Eにしたがって、上位レイヤの画像を同様に配置すると（図8（B））、拡大画像を構成する各画素と、上位レイヤの画像を構成する各画素とは、対応するものどうしが同一の位置に配置されることになる。即ち、この場合、例えば、図8において、上位レイヤの画像（図8（B））である人の部分と、拡大画像（図8（A））の中の人の部分とは、同一の位置に配置されることになる。

【0107】第1および第3のスケラビリティにおける場合も、同様にして、下位レイヤの拡大画像および上位レイヤの画像を構成する、対応する画素どうしが、絶対座標系において同一の位置に配置されるように、オフセットデータFPOS_BおよびFPOS_E、並びにサイズデータFSZ_BおよびFSZ_Eが決定される。

【0108】図3に戻り、画像階層化部21において生成された上位レイヤの画像データ、オフセットデータFPOS_E、およびサイズデータFSZ_Eは、遅延回路22で、後述する下位レイヤ符号化部25における処理時間だけ遅延され、上位レイヤ符号化部23に供給される。また、下位レイヤの画像データ、オフセットデータFPOS_B、およびサイズデータFSZ_Bは、下位レイヤ符号化部25に供給される。また、倍率FRは、遅延回路22を介して、上位レイヤ符号化部23および解像度変換部24に供給される。

【0109】下位レイヤ符号化部25では、下位レイヤの画像データが符号化され、その結果得られる符号化ビットストリームに、オフセットデータFPOS_BおよびサイズデータFSZ_Bが含まれ、多重化部26に供給される。

【0110】また、下位レイヤ符号化部25は、符号化ビットストリームを局所復号し、その結果局所復号結果である下位レイヤの画像データを、解像度変換部24に出力する。解像度変換部24は、下位レイヤ符号化部25からの下位レイヤの画像データを、倍率FRにしたがって拡大（または縮小）することにより、元の大きさに戻し、これにより得られる拡大画像を、上位レイヤ符号化部23に出力する。

【0111】一方、上位レイヤ符号化部23では、上位レイヤの画像データが符号化され、その結果得られる符号化ビットストリームに、オフセットデータFPOS_EおよびサイズデータFSZ_Eが含まれ、多重化部

(15)

特開 2000-13790

27

26に供給される。なお、上位レイヤ符号化部23においては、上位レイヤ画像データの符号化は、後述するように、解像度変換部24から供給される拡大画像をも参照画像として用いて行われる。

【0112】多重化部26では、上位レイヤ符号化部23および下位レイヤ符号化部25の出力が多重化されて出力される。

【0113】なお、下位レイヤ符号化部25から上位レイヤ符号化部23に対しては、下位レイヤのサイズデータFSZ_B、オフセットデータFPOS_B、動きベクトルMV、フラグCODなどが供給されており、上位レイヤ符号化部23では、これらのデータを必要に応じて参照しながら、処理を行うようになされているが、この詳細については、後述する。

【0114】次に、図9は、図3の下位レイヤ符号化部25の詳細構成例を示している。なお、図中、図53における場合と対応する部分については、同一の符号を付してある。即ち、下位レイヤ符号化部25は、基本的には、図53のエンコーダと同様に構成されている。

【0115】画像階層化部21(図3)からの画像データ、即ち、下位レイヤのVOPは、図53における場合と同様に、フレームメモリ31に供給されて記憶され、動きベクトル検出器32において、マクロブロック単位で動きベクトルの検出が行われる。

【0116】但し、下位レイヤ符号化部25の動きベクトル検出器32には、下位レイヤのVOPのサイズデータFSZ_BおよびオフセットデータFPOS_Bが供給されるようになされており、ここでは、このサイズデータFSZ_BおよびオフセットデータFPOS_Bに基づいて、マクロブロックの動きベクトルが検出される。

【0117】即ち、上述したように、VOPは、時刻(フレーム)によって、大きさや位置が変化するため、その動きベクトルの検出にあたっては、その検出のための基準となる座標系を設定し、その座標系における動きを検出する必要がある。そこで、ここでは、動きベクトル検出器32は、上述の絶対座標系を基準となる座標系とし、その絶対座標系に、サイズデータFSZ_BおよびオフセットデータFPOS_Bにしたがって、符号化対象のVOPおよび参照画像とするVOPを配置して、動きベクトルを検出するようになされている。

【0118】なお、検出された動きベクトル(MV)は、予測モードとともに、VLC器36および動き補償器42に供給される他、上位レイヤ符号化部23(図3)にも供給される。

【0119】また、動き補償を行う場合においても、やはり、上述したように、基準となる座標系における動きを検出する必要があるため、動き補償器42には、サイズデータFSZ_BおよびオフセットデータFPOS_Bが供給されるようになされている。

28

【0120】動きベクトルの検出されたVOP(のマクロブロック)は、図53における場合と同様に量子化係数とされてVLC器36に供給される。VLC器36には、やはり図53における場合と同様に、量子化係数、量子化ステップ、動きベクトル、および予測モードが供給される他、画像階層化部21からのサイズデータFSZ_BおよびオフセットデータFPOS_Bも供給されており、ここでは、これらのデータすべてが可変長符号化される。

【0121】動きベクトルの検出されたVOP(のマクロブロック)は、上述したように符号化される他、やはり図53における場合と同様に局所復号され、フレームメモリ41に記憶される。この復号画像は、前述したように参照画像として用いられる他、解像度変換部24(図3)に出力される。

【0122】なお、MPEG4においては、MPEG1および2と異なり、Bピクチャ(B-VOP)も参照画像として用いられるため、Bピクチャも、局所復号され、フレームメモリ41に記憶されるようになされている(但し、現時点においては、Bピクチャが参照画像として用いられるのは上位レイヤについてだけである)。

【0123】一方、VLC器36は、I、P、Bピクチャ(I-VOP、P-VOP、B-VOP)のマクロブロックについて、スキップマクロブロックとするかどうかを決定し、その決定結果を示すフラグCOD、MODBを設定する。このフラグCOD、MODBは、やはり可変長符号化されて伝送される。さらに、フラグCODは、上位レイヤ符号化部23にも供給される。

【0124】次に、図10は、図3の上位レイヤ符号化部23の構成例を示している。なお、図中、図9または図53における場合と対応する部分については、同一の符号を付してある。即ち、上位レイヤ符号化部23は、フレームメモリ52が新たに設けられていることを除けば、基本的には、図9の下位レイヤ符号化部25または図53のエンコーダと同様に構成されている。

【0125】画像階層化部21(図3)からの画像データ、即ち、上位レイヤのVOPは、図53における場合と同様に、フレームメモリ31に供給されて記憶され、動きベクトル検出器32において、マクロブロック単位で動きベクトルの検出が行われる。なお、この場合も、動きベクトル検出器32には、図9における場合と同様に、上位レイヤのVOPの他、そのサイズデータFSZ_EおよびオフセットデータFPOS_Eが供給されるようになされており、動きベクトル検出器32では、上述の場合と同様に、このサイズデータFSZ_EおよびオフセットデータFPOS_Eに基づいて、絶対座標系における上位レイヤのVOPの配置位置が認識され、マクロブロックの動きベクトルが検出される。

【0126】ここで、上位レイヤ符号化部23および下位レイヤ符号化部25における動きベクトル検出器32

(16)

特開2000-13790

30

29

では、図53で説明した場合と同様に、予め設定されている所定のシーケンスにしたがって、VOPが処理されていくが、そのシーケンスは、ここでは、例えば、次のように設定されている。

【0127】即ち、空間スケーラビリティの場合においては、図11(A)または図11(B)に示すように、上位レイヤまたは下位レイヤのVOPは、例えば、P、B、B、B、・・・またはI、P、P、P、・・・ピクチャ(VOP)としてそれぞれ処理されていく。

【0128】そして、この場合、上位レイヤの(表示順で)最初のVOPであるPピクチャ(P-VOP)は、例えば、同時刻における下位レイヤのVOP(ここでは、Iピクチャ(I-VOP))を参照画像として用いて符号化される。また、上位レイヤの2番目以降のVOPであるBピクチャ(B-VOP)は、例えば、その直前の上位レイヤのVOPおよびそれと同時刻の下位レイヤのVOPを参照画像として用いて符号化される。即ち、ここでは、上位レイヤのBピクチャは、下位レイヤのPピクチャと同様に他のVOPを符号化する場合の参照画像として用いられる。

【0129】なお、下位レイヤについては、例えば、MPEG1や2、あるいはH.263における場合と同様に符号化が行われていく。

【0130】SNRスケーラビリティは、空間スケーラビリティにおける倍率FRが1のときと考えられるから、上述の空間スケーラビリティの場合と同様に処理される。

【0131】テンポラルスケーラビリティの場合、即ち、例えば、上述したように、VOが、VOP0、VOP1、VOP2、VOP3、・・・で構成され、VOP1、VOP3、VOP5、VOP7、・・・が上位レイヤとされ(図12(A))、VOP0、VOP2、VOP4、VOP6、・・・が下位レイヤとされた場合においては(図12(B))、図12に示すように、上位レイヤまたは下位レイヤのVOPは、例えば、B、B、B、・・・またはI、P、P、P、・・・ピクチャ(VOP)としてそれぞれ処理されていく。

【0132】そして、この場合、上位レイヤの(表示順で)最初のVOP1(Bピクチャ)は、例えば、下位レイヤのVOP0(Iピクチャ)およびVOP2(Pピクチャ)を参照画像として用いて符号化される。また、上位レイヤの2番目のVOP3(Bピクチャ)は、例えば、その直前にBピクチャとして符号化された上位レイヤのVOP1、およびVOP3の次の時刻(フレーム)における画像である下位レイヤのVOP4(Pピクチャ)を参照画像として用いて符号化される。上位レイヤの3番目のVOP5(Bピクチャ)も、VOP3と同様に、例えば、その直前にBピクチャとして符号化された上位レイヤのVOP3、およびVOP5の次の時刻(フレーム)における画像である下位レイヤのVOP6(P

ピクチャ)を参照画像として用いて符号化される。

【0133】以上のように、あるレイヤのVOP(ここでは、上位レイヤ)については、PおよびBピクチャを符号化するための参照画像として、他のレイヤ(スケーラブルレイヤ)(ここでは、下位レイヤ)のVOPを用いることができる。このように、あるレイヤのVOPを符号化するのに、他のレイヤのVOPを参照画像として用いる場合、即ち、ここでは、上位レイヤのVOPを予測符号化するのに、下位レイヤのVOPを参照画像として用いる場合、上位レイヤ符号化部23(図10)の動きベクトル検出器32は、その旨を示すフラグref_layer_id(階層数が3以上存在する場合、フラグref_layer_idは、参照画像として用いるVOPが属するレイヤを表す)を設定して出力するようになされている。

【0134】さらに、上位レイヤ符号化部23の動きベクトル検出器32は、VOPについてのフラグref_layer_idにしたがい、前方予測符号化または後方予測符号化を、それぞれ、どのレイヤのVOPを参照画像として行うかを示すフラグref_select_code(参照画像情報)を設定して出力するようになされている。

【0135】即ち、例えば、上位レイヤ(Enhancement Layer)のPピクチャが、その直前に復号(局所復号)される、それと同一のレイヤに属するVOPを参照画像として用いて符号化される場合、フラグref_select_codeは「00」とされる。また、Pピクチャが、その直前に表示される、それと異なるレイヤ(ここでは、下位レイヤ)(Reference Layer)に属するVOPを参照画像として用いて符号化される場合、フラグref_select_codeは「01」とされる。さらに、Pピクチャが、その直後に表示される、それと異なるレイヤに属するVOPを参照画像として用いて符号化される場合、フラグref_select_codeは「10」とされる。また、Pピクチャが、それと同時刻における、異なるレイヤのVOPを参照画像として用いて符号化される場合、フラグref_select_codeは「11」とされる。

【0136】一方、例えば、上位レイヤのBピクチャが、それと同時刻における、異なるレイヤのVOPを前方予測のための参照画像として用い、かつ、その直前に復号される、それと同一のレイヤに属するVOPを後方予測のための参照画像として用いて符号化される場合、フラグref_select_codeは「00」とされる。また、上位レイヤのBピクチャが、それと同一のレイヤに属するVOPを前方予測のための参照画像として用い、かつ、その直前に表示される、それと異なるレイヤに属するVOPを後方予測のための参照画像として用いて符号化される場合、フラグref_select_codeは「01」とされる。さらに、上位レイヤの

(17)

特開 2000-13790

31

Bピクチャが、その直前に復号される、それと同一のレイヤに属するVOPを前方予測のための参照画像として用い、かつその直後に表示される、それと異なるレイヤに属するVOPを後方予測のための参照画像として用いて符号化される場合、フラグref_select_codeは「10」とされる。また、上位レイヤのBピクチャが、その直前に表示される、それと異なるレイヤに属するVOPを前方予測のための参照画像として用い、かつその直後に表示される、それと異なるレイヤに属するVOPを後方予測のための参照画像として用いて符号化される場合、フラグref_select_codeは「11」とされる。

【0137】ここで、図11および図12で説明した予測符号化の方法は、1つの例であり、前方予測符号化、後方予測符号化、または両方向予測符号化のための参照画像として、どのレイヤの、どのVOPを用いるかは、例えば、上述した範囲で、自由に設定することが可能である。

【0138】なお、上述の場合においては、便宜的に、「空間スケーラビリティ」、「時間スケーラビリティ」、「SNRスケーラビリティ」という語を用いたが、フラグref_select_codeによって、予測符号化に用いる参照画像を設定する場合、空間スケーラビリティや、テンポラルスケーラビリティ、SNRスケーラビリティを明確に区別することは困難となる。即ち、逆にいえば、フラグref_select_codeを用いることによって、上述のようなスケーラビリティの区別をせずに済むようになる。

【0139】ここで、上述のスケーラビリティとフラグref_select_codeとを対応付けるとすれば、例えば、次のようになる。即ち、Pピクチャについては、フラグref_select_codeが「11」の場合が、フラグref_layer_idが示すレイヤの同時刻におけるVOPを参照画像（前方予測のための参照画像）として用いる場合であるから、これは、空間スケーラビリティまたはSNRスケーラビリティに対応する。そして、フラグref_select_codeが「11」の場合以外は、テンポラルスケーラビリティに対応する。

【0140】また、Bピクチャについては、フラグref_select_codeが「00」の場合が、やはり、フラグref_layer_idが示すレイヤの同時刻におけるVOPを前方予測のための参照画像として用いる場合であるから、これが、空間スケーラビリティまたはSNRスケーラビリティに対応する。そして、フラグref_select_codeが「00」の場合以外は、テンポラルスケーラビリティに対応する。

【0141】なお、上位レイヤのVOPの予測符号化のために、それと異なるレイヤ（ここでは、下位レイヤ）の、同時刻におけるVOPを参照画像として用いる場

32

合、両者の間に動きはないので、動きベクトルは、常に0（0, 0）とされる。

【0142】図10に戻り、上位レイヤ符号化部23の動きベクトル検出器32では、以上のようなフラグref_layer_idおよびref_select_codeが設定され、動き補償器42およびVLC器36に供給される。

【0143】また、動きベクトル検出器32では、フラグref_layer_idおよびref_select_codeにしたがって、フレームメモリ31を参照するだけでなく、必要に応じて、フレームメモリ52をも参照して、動きベクトルが検出される。

【0144】ここで、フレームメモリ52には、解像度変換部24（図3）から、局所復号された下位レイヤの拡大画像が供給されるようになっている。即ち、解像度変換部24では、局所復号された下位レイヤのVOPが、例えば、いわゆる補間フィルタなどによって拡大（補間）され、これにより、そのVOPを、FR倍だけした拡大画像、つまり、その下位レイヤのVOPに対応する上位レイヤのVOPと同一の大きさとした拡大画像が生成され、上位レイヤ符号化部23に供給される。フレームメモリ52では、このようにして解像度変換部24から供給される拡大画像が記憶される。

【0145】従って、倍率FRが1の場合は、解像度変換部24は、下位レイヤ符号化部25からの局所復号されたVOPに対して、特に処理を施すことなく、そのまま、上位レイヤ符号化部23に供給する。

【0146】動きベクトル検出器32には、下位レイヤ符号化部25からサイズデータFSZ_BおよびオフセットデータFPOS_Bが供給されるとともに、遅延回路22（図3）からの倍率FRが供給されるようになされており、動きベクトル検出器32は、フレームメモリ52に記憶された拡大画像を参照画像として用いる場合、即ち、上位レイヤのVOPの予測符号化に、そのVOPと同時刻における下位レイヤのVOPを参照画像として用いる場合（この場合、フラグref_select_codeは、Pピクチャについては「11」に、Bピクチャについては「00」にされる）、その拡大画像に対応するサイズデータFSZ_BおよびオフセットデータFPOS_Bに、倍率FRを乗算する。そして、その乗算結果に基づいて、絶対座標系における拡大画像の位置を認識し、動きベクトルの検出を行う。

【0147】なお、動きベクトル検出器32には、下位レイヤの動きベクトルと予測モードが供給されるようになされており、これは、次のような場合に使用される。即ち、動きベクトル検出器32は、例えば、上位レイヤのBピクチャについてのフラグref_select_codeが「00」である場合において、倍率FRが1であるとき、即ち、SNRスケーラビリティのとき（但し、この場合、上位レイヤの予測符号化に、上位レイヤ

(18)

特開 2000-13790

33

のVOPが用いられるので、この点で、ここでいうSNRスケラビリティは、MPEG2に規定されているものと異なる)、上位レイヤと下位レイヤは同一の画像であるから、上位レイヤのBピクチャの予測符号化には、下位レイヤの同時刻における画像の動きベクトルと予測モードをそのまま用いることができる。そこで、この場合、動きベクトル検出部32は、上位レイヤのBピクチャについては、特に処理を行わず、下位レイヤの動きベクトルと予測モードをそのまま採用する。

【0148】なお、この場合、上位レイヤ符号化部23 10では、動きベクトル検出器32からVLC器36には、動きベクトルおよび予測モードは出力されない(従って、伝送されない)。これは、受信側において、上位レイヤの動きベクトルおよび予測モードを、下位レイヤの復号結果から認識することができるからである。

【0149】以上のように、動きベクトル検出器32は、上位レイヤのVOPの他、拡大画像をも参照画像として用いて、動きベクトルを検出し、さらに、図53で説明したように、予測誤差(あるいは分散)を最小にする予測モードを設定する。また、動きベクトル検出器3 202は、例えば、フラグref_select_codeやref_layer_idその他の必要な情報を設定して出力する。

【0150】なお、図10では、下位レイヤ符号化部25から、下位レイヤにおけるIまたはPピクチャを構成するマクロブロックがスキップマクロブロックであるかどうかを示すフラグCODが、動きベクトル検出器32、VLC器36、および動き補償器42に供給されるようになされている。

【0151】動きベクトルの検出されたマクロブロック 30は、上述した場合と同様に符号化され、これにより、VLC器36からは、その符号化結果としての可変長符号が出力される。

【0152】なお、上位レイヤ符号化部23のVLC器36は、下位レイヤ符号化部25における場合と同様に、フラグCOD、MODBを設定して出力するようになされている。ここで、フラグCODは、上述したように、IまたはPピクチャのマクロブロックがスキップマクロブロックであるかどうかを示すものであるが、フラグMODBは、Bピクチャのマクロブロックがスキップ 40マクロブロックであるかどうかを示すものである。

【0153】また、VLC器36には、量子化係数、量子化ステップ、動きベクトル、および予測モードの他、倍率FR、フラグref_serelect_code、ref_layer_id、サイズデータFSZ_E、オフセットデータFPOS_E、も供給されるようになされており、VLC器36では、これらのデータがすべて可変長符号化されて出力される。

【0154】一方、動きベクトルの検出されたマクロブ 50ロックは符号化された後、やはり上述したように局所復

34

号され、フレームメモリ41に記憶される。そして、動き補償器42において、動きベクトル検出器32における場合と同様にして、フレームメモリ41に記憶された、局所復号された上位レイヤのVOPだけでなく、フレームメモリ52に記憶された、局所復号されて拡大された下位レイヤのVOPをも参照画像として用いて動き補償が行われ、予測画像が生成される。

【0155】即ち、動き補償器42には、動きベクトルおよび予測モードの他、フラグref_serelect_code、ref_layer_id、倍率FR、サイズデータFSZ_B、FSZ_E、オフセットデータFPOS_B、FPOS_Eが供給されるようになされており、動き補償器42は、フラグref_serelect_code、ref_layer_idに基づいて、動き補償すべき参照画像を認識し、さらに、参照画像として、局所復号された上位レイヤのVOP、または拡大画像を用いる場合には、その絶対座標系における位置と大きさを、サイズデータFSZ_EおよびオフセットデータFPOS_E、またはサイズデータFSZ_BおよびオフセットデータFPOS_Bに基づいて認識し、必要に応じて、倍率FRを用いて予測画像を生成する。

【0156】次に、図13は、図1のエンコーダから出力されるビットストリームを復号するデコーダの一実施の形態の構成例を示している。

【0157】このデコーダには、図1のエンコーダから伝送路5または記録媒体6を介して提供される符号化ビットストリームが供給される。即ち、図1のエンコーダから出力され、伝送路5を介して伝送されてくるビットストリームは、図示せぬ受信装置で受信され、あるいは、記録媒体6に記録されたビットストリームは、図示せぬ再生装置で再生され、逆多重化部71に供給される。

【0158】逆多重化部71では、そこに入力された符号化ビットストリーム(後述するVS(Visual Object Sequence))が受信される。さらに、逆多重化部71では、入力されたビットストリームが、VOごとのビットストリームVO#1、VO#2、・・・に分離され、それぞれ、対応するVOP復号部72に供給される。VOP復号部72では、逆多重化部71からのビットストリームから、VOを構成するVOP(画像データ)、サイズデータ(VOP size)、およびオフセットデータ(VOP offset)が復号され、画像再構成部73に供給される。

【0159】画像再構成部73では、VOP復号部72乃至72、それぞれからの出力に基づいて、元の画像が再構成される。この再構成された画像は、例えば、モニタ74に供給されて表示される。

【0160】次に、図14は、スケラビリティを表現する、図13のVOP復号部72の構成例を示している。

(19)

特開 2000-13790

35

【0161】逆多重化部71(図13)から供給されるビットストリームは、逆多重化部91に入力され、そこで、上位レイヤのVOPのビットストリームと、下位レイヤのVOPのビットストリームとに分離される。上位レイヤのVOPのビットストリームは、遅延回路92において、下位レイヤ復号部95における処理の時間だけ遅延された後、上位レイヤ復号部93に供給され、また、下位レイヤのVOPのビットストリームは、下位レイヤ復号部95に供給される。

【0162】下位レイヤ復号部95では、下位レイヤのビットストリームが復号され、その結果得られる下位レイヤの復号画像が解像度変換部94に供給される。また、下位レイヤ復号部95は、下位レイヤのビットストリームを復号することにより得られるサイズデータFSZ__B、オフセットデータFPOS__B、動きベクトル(MV)、予測モード、フラグCODなどの、上位レイヤのVOPを復号するのに必要な情報を、上位レイヤ復号部93に供給する。

【0163】上位レイヤ復号部93では、遅延回路92を介して供給される上位レイヤのビットストリームが、下位レイヤ復号部95および解像度変換部94の出力を必要に応じて参照することにより復号され、その結果得られる上位レイヤの復号画像、サイズデータFSZ__E、およびオフセットデータFPOS__Eが出力される。さらに、上位レイヤ復号部93は、上位レイヤのビットストリームを復号することにより得られる倍率FRを、解像度変換部94に出力する。解像度変換部94では、上位レイヤ復号部93からの倍率FRを用いて、図3における解像度変換部24における場合と同様にし、下位レイヤの復号画像が変換される。この変換により得られる拡大画像は、上位レイヤ復号部93に供給され、上述したように、上位レイヤのビットストリームの復号に用いられる。

【0164】次に、図15は、図14の下位レイヤ復号部95の構成例を示している。なお、図中、図54のデコーダにおける場合と対応する部分については、同一の符号を付してある。即ち、下位レイヤ復号部95は、基本的に、図54のデコーダと同様に構成されている。

【0165】逆多重化部91からの下位レイヤのビットストリームは、バッファ101に供給され、そこで受信されて一時記憶される。IVLC器102は、その後段のブロックの処理状態に対応して、バッファ101からビットストリームを適宜読み出し、そのビットストリームを可変長復号することで、量子化係数、動きベクトル、予測モード、量子化ステップ、サイズデータFSZ__B、オフセットデータFPOS__B、およびフラグCODなどを分離する。量子化係数および量子化ステップは、逆量子化器103に供給され、動きベクトルおよび予測モードは、動き補償器107と上位レイヤ復号部93(図14)に供給される。また、サイズデータFSZ

36

__BおよびオフセットデータFPOS__Bは、動き補償器107、画像再構成部73(図13)、および上位レイヤ復号部93に供給され、フラグCODは、上位レイヤ復号部93に供給される。

【0166】逆量子化器103、IDCT器104、演算器105、フレームメモリ106、または動き補償器107では、図9の下位レイヤ符号化部25の逆量子化器38、IDCT器39、演算器40、フレームメモリ41、または動き補償器42における場合とそれぞれ同様の処理が行われることで、下位レイヤのVOPが復号され、画像再構成部73、上位レイヤ復号部93、および解像度変換部94(図14)に供給される。

【0167】次に、図16は、図14の上位レイヤ復号部93の構成例を示している。なお、図中、図54における場合と対応する部分については、同一の符号を付してある。即ち、上位レイヤ復号部93は、フレームメモリ112が新たに設けられていることを除けば、基本的に、図54のデコーダと同様に構成されている。

【0168】逆多重化部91からの上位レイヤのビットストリームは、バッファ101を介してIVLC器102に供給される。IVLC器102は、上位レイヤのビットストリームを可変長復号することで、量子化係数、動きベクトル、予測モード、量子化ステップ、サイズデータFSZ__E、オフセットデータFPOS__E、倍率FR、フラグref__layer__id、ref__select__code、COD、MODBなどを分離する。量子化係数および量子化ステップは、図15における場合と同様に、逆量子化器103に供給され、動きベクトルおよび予測モードは、動き補償器107に供給される。また、サイズデータFSZ__EおよびオフセットデータFPOS__Eは、動き補償器107および画像再構成部73(図13)に供給され、フラグCOD、MODB、ref__layer__id、およびref__select__codeは、動き補償器107に供給される。さらに、倍率FRは、動き補償器107および解像度変換部94(図14)に供給される。

【0169】なお、動き補償器107には、上述したデータの他、下位レイヤ復号部95(図14)から、下位レイヤの動きベクトル、フラグCOD、サイズデータFSZ__B、およびオフセットデータFPOS__Bが供給されるようになっている。また、フレームメモリ112には、解像度変換部94から拡大画像が供給される。

【0170】逆量子化器103、IDCT器104、演算器105、フレームメモリ106、動き補償器107、またはフレームメモリ112では、図10の上位レイヤ符号化部23の逆量子化器38、IDCT器39、演算器40、フレームメモリ41、動き補償器42、またはフレームメモリ52における場合とそれぞれ同様の処理が行われることで、上位レイヤのVOPが復号され、画像再構成部73に供給される。

(20)

特開 2000-13790

37

【0171】ここで、以上のように構成される上位レイヤ復号部93および下位レイヤ復号部95を有するVOP復号部72.においては、上位レイヤについての復号画像、サイズデータFSZ_E、およびオフセットデータFPOS_E（以下、適宜、これらをすべて含めて、上位レイヤデータという）と、下位レイヤについての上位レイヤについての復号画像、サイズデータFSZ_B、およびオフセットデータFPOS_B（以下、適宜、これらをすべて含めて、下位レイヤデータという）が得られるが、画像再構成部73では、この上位レイヤデータまたは下位レイヤデータから、例えば、次のようにして画像が再構成されるようになされている。

【0172】即ち、例えば、第1の空間スケーラビリティ（図4）が行われた場合（入力されたVOP全体が上位レイヤとされるときに、そのVOP全体を縮小したものが下位レイヤとされた場合）において、下位レイヤデータおよび上位レイヤデータの両方のデータが復号されたときには、画像再構成部73は、上位レイヤデータのみに基づき、サイズデータFSZ_Eに対応する大きさの上位レイヤの復号画像（VOP）を、オフセットデータFPOS_Eによって示される位置に配置する。また、例えば、上位レイヤのビットストリームにエラーが生じたり、また、モニター4が、低解像度の画像にしか対応していないため、下位レイヤデータのみの復号が行われたときには、画像再構成部73は、その下位レイヤデータのみに基づき、サイズデータFSZ_Bに対応する大きさの上位レイヤの復号画像（VOP）を、オフセットデータFPOS_Bによって示される位置に配置する。

【0173】また、例えば、第2の空間スケーラビリティ（図5）が行われた場合（入力されたVOPの一部が上位レイヤとされるときに、そのVOP全体を縮小したものが下位レイヤとされた場合）において、下位レイヤデータおよび上位レイヤデータの両方のデータが復号されたときには、画像再構成部73は、サイズデータFSZ_Bに対応する大きさの下位レイヤの復号画像を、倍率FRにしたがって拡大し、その拡大画像を生成する。さらに、画像再構成部73は、オフセットデータFPOS_BをFR倍し、その結果得られる値に対応する位置に、拡大画像を配置する。そして、画像再構成部73は、サイズデータFSZ_Eに対応する大きさの上位レイヤの復号画像を、オフセットデータFPOS_Eによって示される位置に配置する。

【0174】この場合、上位レイヤの復号画像の部分が、それ以外の部分に比較して高い解像度で表示されることになる。

【0175】なお、上位レイヤの復号画像を配置する場合においては、その復号画像と、拡大画像とは合成される。

【0176】また、図14（図13）には図示しなかつ

38

たが、上位レイヤ復号部93（VOP復号部72.）から画像再構成部73に対しては、上述したデータの他、倍率FRも供給されるようになされており、画像再構成部73は、これを用いて、拡大画像を生成するようになされている。

【0177】一方、第2の空間スケーラビリティが行われた場合において、下位レイヤデータのみが復号されたときには、上述の第1の空間スケーラビリティが行われた場合と同様にして、画像が再構成される。

【0178】さらに、第3の空間スケーラビリティ（図6、図7）が行われた場合（入力されたVOPを構成する物体ごとに、その物体（オブジェクト）全体を上位レイヤとするとともに、その物体全体を囲引いたものを下位レイヤとした場合）においては、上述の第2の空間スケーラビリティが行われた場合と同様にして、画像が再構成される。

【0179】上述したように、オフセットデータFPOS_BおよびFPOS_Eは、下位レイヤの拡大画像および上位レイヤの画像を構成する、対応する画素どうし、絶対座標系において同一の位置に配置されるようになっているため、以上のように画像を再構成することで、正確な（位置ずれのない）画像を得ることができ

る。

【0180】次に、図1のエンコーダが出力する符号化ビットストリームのシンタクスについて、例えば、MPEG4規格のVideo Verification Model (Version10.0)（以下、適宜、VM10.0と記述する）を例に説明する。

【0181】図17は、VM10.0における符号化ビットストリームの構成を示している。

【0182】符号化ビットストリームは、VS (Visual Object Sequence) を単位として構成される。ここで、VSは、画像シーケンスであり、例えば、一本の番組や映画などに相当する。

【0183】各VSは、1以上のVISO (Visual Object) から構成される。ここで、VISOには、幾つかの種類がある。即ち、VISOには、例えば、静止画であるステルテクスチャオブジェクト (Still Texture Object) や、顔画像から構成されるフェイスオブジェクト (Face Object)、動画像のオブジェクトであるVO (Video Object) などがある。従って、符号化ビットストリームが動画像のものである場合、VISOは、VOから構成される。

【0184】VOは、1以上のVOL (Video Object Layer) から構成される（画像を階層化しないときは1のVOLで構成され、画像を階層化する場合には、その階層数だけのVOLで構成される）。

【0185】VOLは、必要な数のGOV (Group of VOP) で構成され、GOVは、1以上のVOP (Video Object Plane) のシーケンスで構成される。なお、GOVはなくても良く、この場合、VOLは、1以上のVOP

で構成されることになる。

【0186】図18または図19は、VSまたはVOのシンタクスをそれぞれ示している。VOは、画像全体または画像の一部（物体）のシーケンスに対応するビットストリームであり、従って、VSは、そのようなシーケンスの集合で構成される（よって、VSは、例えば、一本の番組などに相当する）。

【0187】図20乃至図25は、VOLのシンタクスを示している。

【0188】VOLは、上述したようなスケーラビリティのためのクラスであり、video_object_layer_idで示される番号によって識別される。即ち、例えば、下位レイヤのVOLについてのvideo_object_layer_idは0とされ、また、例えば、上位レイヤのVOLについてのvideo_object_layer_idは1とされる。なお、上述したように、スケーラブルのレイヤの数は2に限られることなく、1や3以上を含む任意の数とすることができる。

【0189】また、各VOLについて、それが画像全体であるのか、画像の一部であるのかは、video_object_layer_shapeで識別される。このvideo_object_layer_shapeは、VOLの形状を示すフラグで、例えば、以下のよう

に設定される。

【0190】即ち、VOLの形状が長方形状であるとき、video_object_layer_shapeは、例えば「00」とされる。また、VOLが、ハードキー（0または1のうちのいずれか一方の値をとる2値（Binary）の信号）によって抜き出される領域の形状をしているとき、video_object_layer_shapeは、例えば「01」とされる。さらに、VOLが、ソフトキー（0乃至1の範囲の連続した値（Gray-Scale）をとることが可能な信号（MP EG 4では、8ビットで表現される））によって抜き出される領域の形状をしているとき（ソフトキーを用いて合成されるものであるとき）、video_object_layer_shapeは、例えば「10」とされる。

【0191】ここで、video_object_layer_shapeが「00」とされるのは、VOLの形状が長方形状であり、かつ、そのVOLの絶対座標系における位置および大きさが、時間とともに変化しない、即ち、一定の場合である。なお、この場合、その大きさ（横の長さとの縦の長さ）は、video_object_layer_widthとvideo_object_layer_heightによって示される。video_object_layer_widthおよびvideo_object_layer_heightは、いずれも10ビットの固定長のフラグで、video_object_layer_shapeが「00」の場合には、最初に、一度だけ伝送される（これは、video_object_layer_shapeが「00」の場合、上述したように、VOLの絶対座標系における大きさが一定であるからである）。

【0192】また、VOLが、下位レイヤまたは上位レイヤのうちのいずれであるかは、1ビットのフラグであるscalabilityによって示される。VOLが下位レイヤ

の場合、scalabilityは、例えば1とされ、それ以外の場合、scalabilityは、例えば0とされる。

【0193】さらに、VOLが、自身以外のVOLにおける画像を参照画像として用いる場合、その参照画像が属するVOLは、上述したように、ref_layer_idで表される。なお、ref_layer_idは、上位レイヤについてのみ伝送される。

【0194】また、VOLにおいて、hor_sampling_factor_nとhor_sampling_factor_mは、下位レイヤのVOPの水平方向の長さに対応する値と、上位レイヤのVOPの水平方向の長さに対応する値をそれぞれ示す。従って、下位レイヤに対する上位レイヤの水平方向の長さ（水平方向の解像度の倍率）は、式hor_sampling_factor_n/hor_sampling_factor_mで与えられる。

【0195】さらに、VOLにおいて、ver_sampling_factor_nとver_sampling_factor_mは、下位レイヤのVOPの垂直方向の長さに対応する値と、上位レイヤのVOPの垂直方向の長さに対応する値をそれぞれ示す。従って、下位レイヤに対する上位レイヤの垂直方向の長さ（垂直方向の解像度の倍率）は、式ver_sampling_factor_n/ver_sampling_factor_mで与えられる。

【0196】図26および図27は、GOVのシンタクス（Syntax）を示している。

【0197】GOV層は、符号化ビットストリームの先頭だけでなく、符号化ビットストリームの任意の位置に挿入することができるように、VOL層とVOP層との間に規定されている（図17）。これにより、あるVOL#0が、VOP#0, VOP#1, ..., VOP#n, VOP#(n+1), ..., VOP#mといったVOPのシーケンスで構成される場合において、GOV層は、その先頭のVOP#0の直前だけでなく、VOP#(n+1)の直前（VOP#nとVOP#(n+1)との間）にも挿入することができる。従って、エンコードにおいて、GOV層は、例えば、符号化ストリームの中の、ランダムアクセスさせたい位置に挿入することができ、このGOV層を挿入することで、あるVOLを構成するVOPの一連のシーケンスは、複数のグループ（GOV）に分けられて符号化されることになる。

【0198】GOV層は、図26に示すように、グループスタートコード（group_start_code）、VOPタイムインクリメントレゾリューション（VOP_time_increment_resolution）、タイムコード（time_code）、クローズドGOP（closed_gop）、ブローケンリンク（broken_link）、GOVを構成するVOP（Group_of_VideoObjectPlane()）、ネクストスタートコード（next_start_code()）が順次配置されて構成される。

【0199】次に、GOV層のセマンティクス（Semantics）について説明するが、GOV層のセマンティクスは、基本的には、MP EG 2のGOP層と同様であり、従って、特に記述しない部分については、MP EG 2 Video規格（ISO/IEC 13818-2）を参照されたい。

【0200】まず、group_start_codeは、000001B8 (16進数)で、GOVの開始位置を示す。VOP_time_increment_resolutionは、後述するVOP_time_incrementによって示される同期点どうしの間隔である1秒間を、何分割するかで分割数(分解能)を表す。

【0201】time_codeは、図27に示すように、時刻の時間の単位を表すtime_code_hours、時刻の分の単位を表すtime_code_minutes、marker_bit、および時刻の秒の単位を表すtime_code_secondsで構成される。そして、このtime_codeのうちの、time_code_hours、time_code_minutes、time_code_secondsによって、GOVの先頭の時刻が表される。その結果、GOV層のtime_code(秒精度先頭表示時刻)は、秒精度で、その先頭の時刻、即ち、そのGOV層の符号化が開始された、VOPのシーケンス上の絶対時刻を表現することとなる。具体的には、time_codeには、GOVの中で最初に表示されるVOPの絶対的な表示時刻を秒精度で表したものが設定される。

【0202】なお、VM10.0において、time_codeは、「The parameters correspond to those defined in the IEC standard publication 461 for "time and control codes for video tape recorders". The timecode refers to the first plane (in display order) after the GOV header」と記載されている。

【0203】time_codeのmarker_bitは、符号化ビットストリームにおいて、0が23個以上連続しないように1とされる。

【0204】next_start_codeは、次のGOVの先頭の位置を与える。

【0205】以上のようなGOVによれば、そのタイムコードtime_codeにより、GOVのヘッダに続く、表示順で、最初に表示されるVOPの絶対的な表示時刻を秒精度で認識することが可能となる。なお、上述のように、GOV層のtime_codeは秒精度であるため、VOPの表示時刻の、さらに細かい精度の部分は、VOP毎に設定されるが、この点については、後述する。

【0206】なお、GOV層は、前述したように、本件出願人による提案後に、MPEG4において導入されたものである。

【0207】次に、図28乃至図36は、VOP (Video Object Plane Class) のシンタクスを示している。

【0208】VOPの大きさ(横と縦の長さ)は、例えば、10ビット固定長のVOP_widthとVOP_heightで表される。また、VOPの絶対座標系における位置は、例えば、10ビット固定長のVOP_horizontal_spatial_mc_refとVOP_vertical_mc_refで表される。なお、VOP_widthまたはVOP_heightは、VOPの水平方向または垂直方向の長さをそれぞれ表し、これらは、上述のサイズデータFSZ__BやFSZ__Eに相当する。また、VOP_horizontal_spatial_mc_refまたはVOP_vertical_mc_refは、VOPの水平方向または垂直方向の座標(xまたはy座

標)をそれぞれ表し、これらは、上述のオフセットデータFPOS__BやFPOS__Eに相当する。

【0209】VOP_width、VOP_height、VOP_horizontal_spatial_mc_ref、およびVOP_vertical_mc_refは、video_object_layer_shapeが「00」以外の場合にのみ伝送される。即ち、video_object_layer_shapeが「00」の場合、上述したように、VOPの大きさおよび位置はいずれも一定であるから、VOP_width、VOP_height、VOP_horizontal_spatial_mc_ref、およびVOP_vertical_mc_refは伝送する必要がない。この場合、受信側では、VOPは、その左上の頂点が、例えば、絶対座標系の原点に一致するように配置され、また、その大きさは、図20乃至図25に示したVOLのvideo_object_layer_widthおよびvideo_object_layer_heightから認識される。

【0210】なお、VOPにおいて、ref_select_codeは、上述したように、参照画像として用いる画像を表す。

【0211】ところで、VM10.0では、各VOP (Video Object Plane: 従来のFrameに相当する) の表示時刻は、そのVOPで規定されているmodulo_time_baseおよびVOP_time_increment (図28)、並びにそのVOPから構成されるGOVで規定されているtime_code (図27) によって、次のように定められる。

【0212】即ち、modulo_time_baseは、エンコーダのローカルな時間軸における時刻を、1秒(1000ms(ミリ秒))の精度で表す。modulo_time_baseは、VOPヘッダの中で伝送されるマーカ(marker)で表現され、必要な数の「1」と、1つの「0」とで構成される。modulo_time_baseを構成する「1」の数が、最後に(現在から遡って、最も最近に)(直前に)表示されたI-VOP、P-VOPで符号化/復号化されたmodulo_time_base、またはGOVヘッダのtime_codeによって示された同期点(1秒精度の時刻)からの累積時間を表す。

【0213】具体的には、VOPのmodulo_time_baseが、例えば、「0」の場合は、直前に表示されたI-VOP、P-VOPのmodulo_time_base、またはGOVヘッダのtime_codeによって示された同期点からの累積時間が0秒であることを表す。また、modulo_time_baseが、例えば、「10」の場合は、直前に表示されたI-VOP、P-VOPのmodulo_time_base、またはGOVヘッダのtime_codeによって示された同期点からの累積時間が1秒であることを表す。さらに、modulo_time_baseが、例えば、「110」の場合は、直前に表示されたI-VOP、P-VOPのmodulo_time_base、またはGOVヘッダのtime_codeによって示された同期点からの累積時間が2秒であることを表す。以上のように、VOPのmodulo_time_baseの「1」の数は、その直前に表示されたI-VOP、P-VOPのmodulo_time_base、またはGOVヘッダのtime_codeによって示された同期点からの秒数になっている。

(23)

特開 2000-13790

43

【0214】なお、VM10.0では、modulo_time_baseについて、「This value represents the local time base at the one second resolution unit (1000 milliseconds). It is represented as a marker transmitted in the VOP header. The number of consecutive "1" followed by a "0" indicates the number of seconds has elapsed since the synchronization point marked by the modulo_time_base of the last displayed I/P-VOPs belonging to the same VOL. There are two exceptions, one for the first I/P-VOP after the GOV header, and the other is for B-VOPs (in display order) to the first I-VOP after the GOV header. For the first I/P-VOP after the GOV header, the modulo_time_base indicates the time relative to the time_code in the GOV header. For the B-VOPs prior (in display order) to the first I-VOP after the GOV header, the modulo_time_base indicates the time relative to the time_code in the GOV header」と記載されている。

【0215】即ち、ある注目VOPのmodulo_time_baseは、直前に表示されたI-VOP、P-VOPのmodulo_time_baseによって示される同期点、即ち、直前に表示されたI-VOP、P-VOPの表示時刻の秒精度の時刻からの相対時間によって、注目VOPの秒精度の表示時刻を表す。但し、GOVヘッダに続いて最初に符号化／復号されるI-VOPまたはP-VOPのmodulo_time_baseは、GOVヘッダのtime_codeからの相対時間によって、そのI-VOPまたはP-VOPの秒精度の表示時刻を表し、また、GOVヘッダに続いて最初に符号化／復号されるI-VOPより前に表示されるB-VOPのmodulo_time_baseも、GOVヘッダのtime_codeからの相対時間によって、そのB-VOPの秒精度の表示時刻を表す。

【0216】なお、GOVヘッダに続いて最初に符号化／復号されるI-VOPまたはP-VOP、およびそのようなI-VOPより前に表示されるB-VOPに関して、GOVヘッダのtime_codeによって示される同期点を、直前に表示されたI-VOP、P-VOPのmodulo_time_baseによって示された同期点（直前に表示されたI-VOP、P-VOPの表示時刻の秒精度の時刻）と考えれば、VOPのmodulo_time_baseは、どのようなVOPについても、直前に表示されたI-VOP、P-VOPのmodulo_time_baseによって示された同期点からの時間によって、そのVOPの表示時刻を、秒精度で表すことができる。

【0217】VOP_time_incrementは、1秒以下の精度でVOPの表示時刻を表すのに使用される。すなわちtime_code及びmodulo_time_baseは、1秒の精度で時刻を表すが、それ以下の精度は、VOP_time_incrementを用いて表現される。VM10.0の場合、VOP_time_incrementの精度は、上述したGOVレイヤのtime_increment_resolution

44

(図26)で示され、VOP_time_increment_resolution = 1000とすると、VOP_time_incrementは、VOPの表示時刻を、1ms (= 1000ミリ秒/VOP_time_increment_resolution) 単位で表すものとなる。

【0218】図37および図38は、以上の定義に基づいて、time_code、modulo_time_baseとVOP_time_incrementとの関係を示した図である。

【0219】図37において、VOは、GOVヘッダ、B₁(B-VOP)、I₁(I-VOP)、B₂、B₃、P₁(P-VOP)....というVOPのシーケンスで構成されている。いま、GOVヘッダのtime_codeが時刻t₀ = 0h:12m:43Sec (0時間12分43秒)を示していたとすると、modulo_time_baseは、time_codeを基準とする時刻を、1秒精度で表し、従って、t₀+1秒、t₀+2秒、...という時刻を、同期点として表す。なお、図37において、表示順は、B₁、I₁、B₂、B₃、P₁、...であるが、符号化／復号順は、GOVヘッダ、I₁、B₁、P₁、B₂、B₃、...である。

【0220】図37では、(後述する図38、図46、図47においても同様)、各VOPについてのVOP_time_incrementを、四角形で囲んだ数字で、modulo_time_baseを、四角形とダブルクォーテーション(") で囲んだビット列で、それぞれ示してあり、さらに、VOP_time_increment_resolutionを1000としてある。従って、図37では、B₁、I₁、B₂、B₃、P₁についてのVOP_time_incrementがそれぞれ350、750、150、550、350とされており、また、VOP_time_increment_resolutionが1000であるから、各VOPの表示時刻の1秒以下の値は、それぞれ350ms (350/1000秒)、750ms (750/1000秒)、150ms (150/1000秒)、550ms (550/1000秒)、350ms (350/1000秒)となる。

【0221】いま、図37において、B₁は、GOVヘッダが符号化／復号された後、最初に符号化／復号されるI-VOPであるI₁より前に表示されるB-VOPであるから、その表示時刻の基準は、GOVヘッダのtime_codeで示される同期点である時刻t₀となる。いま、B₁のtime_modulo_baseは、"0"であり、従って、B₁の表示時刻は、GOVヘッダのtime_codeで示された時刻t₀の、0+350/1000秒後の時刻、すなわち0h:12m:34s+0s:350ms=0h:12m:34s:350msとなる。

【0222】次に、I₁は、GOVヘッダが符号化／復号された後、最初に符号化／復号されるI-VOPであるから、その表示時刻の基準は、GOVヘッダのtime_codeで示される同期点である時刻t₀となる。いま、I₁のtime_modulo_baseは、"0"であり、従って、I₁の表示時刻は、GOVヘッダのtime_codeで示された時刻t₀の、0+750/1000秒の時刻、すなわち0h:12m:34s+0s:750ms=0h:12m:34s:750msとなる。

【0223】次に、B₂は、I₁の直後に表示されるた

(24)

特開 2000-13790

45

め、その表示時刻の基準は、 I_1 の modulo_time_base によって示された同期点となる。即ち、 I_1 の表示時刻は、上述したように、0h:12m:34s:750msであり、modulo_time_base は、時刻 $t_1 = 0h:12m:34s$ から 0 秒後の時刻である 0h:12m:34s を、同期点として示している。そして、 B_1 の modulo_time_base は、"10" であることから、 B_1 の表示時刻は、 I_1 の modulo_time_base によって示される同期点の、 $1 + 150 / 1000$ 秒後の時刻、すなわち 0h:12m:34s+1s:150ms=0h:12m:35s:150ms となる。 B_1 の表示時刻も、 B_1 の表示時刻と同様に求めることができ、 I_2 の modulo_time_base によって示される同期点の、 $1 + 550 / 1000$ 秒後の時刻、すなわち 0h:12m:34s+1s:550ms=0h:12m:35s:550ms となる。

【0224】次に、 P_1 については、その前に表示される B_1 、 B_1 は共に B-VOP であるため、これらの modulo_time_base によって示される同期点は、 P_1 の表示時刻の計算には用いられず、 P_1 の直前に表示される I/P-VOP である I_1 の modulo_time_base によって示される同期点 (0h:12m:34s) が用いられる。また、図 37 では、 P_1 の time_modulo_base が、"110" であり、time_increment は、上述したように 350 であることから、その表示時刻は、 I_1 の modulo_time_base によって示される同期点の、 $2 + 350 / 1000$ 秒後の時刻、すなわち 0h:12m:34s+2s:350ms=0h:12m:36s:350ms となる。

【0225】次に、図 38 では、VO が、GOV ヘッダ B_1 、 B_1 、 I_1 、 B_1 、 P_1 、... という VOP のシーケンスで構成されている。また、GOV ヘッダの time_code は、図 37 における場合と同様に、時刻 $t_1 = 0h:12m:43s$ (0 時間 12 分 43 秒) を示している。

【0226】図 38 において、 B_1 については、modulo_time_base が "0" と、VOP_time_increment が 350 となっている。そして、 B_1 は、GOV ヘッダが符号化/復号された後、最初に符号化/復号される I-VOP である I_1 より前に表示される B-VOP であるから、その表示時刻は、GOV ヘッダの time_code で示される同期点である時刻 t_1 を基準に計算される。従って、 B_1 の表示時刻は、図 37 の B_1 と同様に、0h:12m:34s+0s:350ms=0h:12m:34s:350ms となる。

【0227】次に、 B_2 は、 B_1 と同様に、GOV ヘッダが符号化/復号された後、最初に符号化/復号される I-VOP である I_1 より前に表示される B-VOP であるから、その表示時刻は、やはり、GOV ヘッダの time_code で示される同期点である時刻 t_1 を基準に計算される。いま、 B_2 の time_modulo_base が "0" で、VOP_time_increment が 750 となっているから、 B_2 の表示時刻は、GOV ヘッダの time_code で示された時刻 t_1 の、 $0 + 750 / 1000$ 秒後の時刻、すなわち 0h:12m:34s+0s:750ms=0h:12m:34s:750ms となる。

【0228】次に、 I_1 は、GOV の符号化/復号後、最初に表示される I-VOP であり、その modulo_time_

46

base が "01" で、VOP_time_increment が 150 であるため、その表示時刻は、time_code によって示された同期点より $1 + 150 / 1000$ 秒遅れとなる。従って、 I_1 の表示時刻は、時刻 t_1 の、 $0 + 750 / 1000$ 秒後の時刻、すなわち 0h:12m:34s+1s:150ms=0h:12m:35s:150ms となる。

【0229】次に、 B_2 については、その直前に表示される I/P-VOP である I_1 の modulo_time_base によって示される同期点 (I_1 の表示時刻の秒精度)、即ち、0h:12m:35s を基準に、その表示時刻が計算される。いま、 B_2 の modulo_time_base は "0" で、VOP_time_increment が 550 であるため、その表示時刻は、0h:12m:35s+0s:550ms=0h:12m:35s:550ms となる。

【0230】次に、 P_1 については、その直前に表示される I/P-VOP である I_1 の modulo_time_base によって示される同期点、即ち、 B_1 と同様に、0h:12m:35s を基準に、その表示時刻が計算される。いま、 P_1 の modulo_time_base は "10" で、VOP_time_increment が 350 であるため、その表示時刻は、0h:12m:35s+1s:350ms=0h:12m:36s:350ms となる。

【0231】ここで、以上のように、VM10.0 では、表示時刻の計算に、I/P-VOP の modulo_time_base によって示される同期点 (これは、その表示時刻の秒精度に等しい) は用いられるが、B-VOP の modulo_time_base によって示される同期点は用いられない。これは、主として次のような理由による。即ち、B-VOP は、表示順では、I/P-VOP に挟まれているが、符号化/復号化順では、B-VOP を挟んでいる I/PVOP が符号化/復号された後に符号化/復号される。このため、仮に、表示時刻の計算に、B-VOP の modulo_time_base によって示される同期点を用いることとすると、即ち、B-VOP の表示時刻の秒精度を用いることとすると、表示順において B-VOP の直後に位置する I/P-VOP は、すでに符号化/復号されているにもかかわらず、その B-VOP の符号化/復号が終了するまでは、I/P-VOP の表示時刻を求めることができず、処理が煩雑になるためである。

【0232】次に、VM10.0 では、上述したとおり、GOV レイヤを採用しており、これにより、効率的なランダムアクセスを可能としている。即ち、GOV レイヤの time_code により、GOV の絶対的な開始時刻 (但し、秒精度) が定まり、この開始時刻と time_modulo_base、VOP_time_increment により、上述したように、その GOV に含まれる各 VOP の絶対的な表示時刻が容易に求められる。そして、この表示時刻を参照することにより、符号化ビットストリームの途中にある VOP にアクセスすること、すなわちランダムアクセスすることが可能となる。

【0233】ところで、VM10.0 では、"The GOV layer is an optional layer, so the bitstream can have any (include non) number of the GOV header, and the frequency of the GOV header is an encoder issue. Sin

ce the GOV header shall be followed by the I-VO P.」と定義されていることから、GOVレイヤの挿入については、符号化ビットストリームの中で、GOVヘッダの後にI-VOPを配置しなければならないという制限がある点を除き、エンコーダ側で、その数と頻度を、自由に設定することができる。

【0234】即ち、VMI0.0では、スケーラビリティを実現するための階層化を行った場合に、各階層におけるGOVの挿入位置は、特に規定されていない。従って、例えば、上位レイヤと下位レイヤとの2階層に階層化した場合には、上位レイヤまたは下位レイヤそれぞれについて、独立に、I-VOPを、ヘッダの後に配置（但し、符号化ビットストリームの中で）した形のGOVを構成することができる。

【0235】一方、MPEG4では、上位レイヤは、必ずしもI-VOPを使用しなくてもエンコード／デコードが可能である。これは、上位レイヤについては、基本的に、下位レイヤのVOPが参照画像として用いられるからであり、MPEG4では、このように、下位レイヤのVOPを参照画像として用いることで、符号化効率を向上させている。

【0236】しかしながら、上述したように、GOVを用いるためには、符号化／復号化順でGOVヘッダの後にI-VOPを配置しなければならない。そして、I-VOPのデータ量は、一般に、P-VOPやB-VOPのデータ量より大きいから、I-VOPを使用しなくても問題のない上位レイヤにI-VOPを使用しなくてはならないことは、符号化効率を低下させることになる。

【0237】そこで、符号化効率の低下を防止するために、VM10.0で規定されている上述の定義は、「The GOV layer is an optional layer, so the bitstream can have any (include non) number of the GOV header, and the frequency of the GOV header is an encoder issue. Since the GOV header shall be followed by the I-VOP in base layer.」と変更し、下位レイヤについてのみ、GOVヘッダの後にI-VOPを配置しなければならないという制限をつけるべきであり、上位レイヤについては、そのような制限をつけるべきではない。

【0238】図1のエンコーダまたは図13のデコーダは、上述の変更後の定義にしたがって、それぞれ符号化または復号が行われるようになされている。

【0239】次に、VM10.0では、下位レイヤと上位レイヤのGOVの時間的な位置関係については、特に制限がない。従って、下位レイヤのGOVの最初のVOPまたは最後のVOPそれぞれと、上位レイヤの最初のVOPまたは最後のVOPそれぞれとの時間的な位置は、必ずしも一致しなくてもよいこととなる。

【0240】しかしながら、下位レイヤと上位レイヤとで、GOVの挿入位置が異なる場合に、その符号化ビットストリームにランダムアクセスするときには、上位レイ

ヤのVOPが参照する下位レイヤの復号画像を得るために、本来表示されない下位レイヤのVOPを復号しなければならないケースが生じ、この場合、迅速で、効率的なランダムアクセスが困難となる。

【0241】即ち、いま、下位レイヤの、表示順で# i 番目のI-VOPまたはP-VOPを、それぞれ、 I_{i1} 、または P_{i1} とそれぞれ表すとともに、上位レイヤの、表示順で# i 番目のP-VOPを、 P_{i1} と表し、例えば、図39に示すように、 I_{i1} 、 P_{i2} 、 P_{i3} 、 P_{i4} 、 P_{i5} 、 I_{i6} 、 P_{i7} 、 P_{i8} 、 P_{i9} 、 P_{i10} 、 I_{i11} 、 P_{i12} 、 P_{i13} 、・・・で構成される下位レイヤと、 P_{i1} 、 P_{i2} 、 P_{i3} 、 P_{i4} 、 P_{i5} 、 P_{i6} 、 P_{i7} 、 P_{i8} 、 P_{i9} 、 P_{i10} 、 P_{i11} 、 P_{i12} 、 P_{i13} 、・・・で構成される上位レイヤとからなる符号化ビットストリームを考える。

【0242】図39において、下位レイヤの $I_{0,1}$, $P_{0,1}$, $P_{0,2}$, $P_{0,3}$, $P_{0,4}$, $I_{0,5}$, $P_{0,5}$, $P_{0,6}$, $P_{0,7}$, $P_{0,8}$, $P_{0,9}$, $P_{0,10}$, $I_{0,11}$, $P_{0,12}$, $P_{0,13}$, ...それぞれと、上位レイヤの $P_{1,1}$, $P_{1,2}$, $P_{1,3}$, $P_{1,4}$, $P_{1,5}$, $P_{1,6}$, $P_{1,7}$, $P_{1,8}$, $P_{1,9}$, $P_{1,10}$, $P_{1,11}$, $P_{1,12}$, $P_{1,13}$, ...それぞれとは、同一時刻に表示されるVOPであり、さらに、上位レイヤの $P_{1,1}$ 乃至 $P_{1,13}$ それぞれは、下位レイヤの $I_{0,1}$ 乃至 $P_{0,13}$ それぞれを参照画像として符号化されている。即ち、空間スケラビリティが実現されている。そして、図39では、下位レイヤの $I_{0,1}$ 乃至 $P_{0,4}$ がGOV_{0,1}を、 $I_{0,5}$ 乃至 $P_{0,8}$ がGOV_{0,2}を、それぞれ構成しており、上位レイヤの $P_{1,1}$ 乃至 $P_{1,4}$ がGOV_{1,1}を、 $P_{1,5}$ 乃至 $P_{1,8}$ がGOV_{1,2}を、 $P_{1,9}$ 乃至 $P_{1,12}$ がGOV_{1,3}を、それぞれ構成している。即ち、下位レイヤでは、5つのVOPごとにGOVが構成され、上位レイヤでは、4つのVOPごとにGOVが構成されている。

【0243】この場合において、例えば、上位レイヤのGOV_iの先頭にランダムアクセスするとしたとき、その先頭のP_iは、下位レイヤのGOV_jのP_jの復号画像を参照画像として用いて復号されるから、P_iを復号する必要がある。しかしながら、P_jは、下位レイヤのGOV_jを構成するVOP（最後に表示されるVOP）であり、従って、P_jを復号するには、GOV_jを構成するVOPを順次復号する必要がある。そして、P_iが復号されるまでの間は、そのP_jを参照画像として用いる上位レイヤのP_iを復号することができないから、迅速なランダムアクセスが困難となる。

【0244】このようなことは、時間スケールビリティの場合も、同様に生ずる。

【0245】即ち、いま、図40に示すように、 b_1 , b_2 , b_3 , b_4 , b_5 , \dots で表されるVOPで構成される下位レイヤと、 e_1 , e_2 , e_3 , e_4 , e_5 , e_6 , e_7 , e_8 , \dots で表されるVOPで構成される上位レイヤとからなる符号化ビットストリームを考える。なお、 e_1 , b_1 , e_1 , b_2 , e_2 , e_1 , b_3 , e_3 , e_1 , b_4 , e_4 , e_1 , b_5 , \dots が、一連の画像シーケンスを構

(26)

特開 2000-13790

49

成しており、図40では、そのような画像シーケンスから、 $b_1, b_2, b_3, b_4, b_5, \dots$ が抽出されて下位レイヤが構成され、 $e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, \dots$ が抽出されて上位レイヤが構成されている。

【0246】そして、図40では、下位レイヤの b_1 乃至 b_5 がGOV₁を、 b_6 乃至 b_8 がGOV₂を、それぞれ構成しており、上位レイヤの e_1 乃至 e_8 がGOV₁を、 e_9 乃至 e_{12} がGOV₂を、 e_{13} 乃至 e_{16} がGOV₃を、それぞれ構成している。即ち、下位レイヤおよび上位レイヤの両方とも、3つのVOPごとにGOVが構成されている。なお、 e_1 は b_1 を、 e_2 および e_3 は b_2 を、 e_4 および e_5 は b_3 を、 e_6 および e_7 は b_4 を、 e_8 および e_9 は b_5 を、それぞれ参照画像として用いている。

【0247】この場合において、例えば、上位レイヤのGOV₁の先頭にランダムアクセスするとしたとき、その先頭の e_1 は、下位レイヤのGOV₁の b_1 の復号画像を参照画像として用いて復号されるから、まずは、 b_1 を復号する必要がある。しかしながら、 b_1 は、下位レイヤのGOV₁を構成するVOP（最後に表示されるVOP）であり、従って、 e_1 を復号するには、GOV₁を構成するVOPを順次復号する必要がある。そして、 b_1 を復号するまでの間は、ランダムアクセスする目的である上位レイヤの e_1 を復号することができないから、やはり、迅速なランダムアクセスが困難となる。

【0248】そこで、下位レイヤのGOVの最初に表示されるVOPの表示時刻と同時刻またはその直後に表示される上位レイヤのVOPが、上位レイヤのGOVの最初に表示されるものとなるように、上位レイヤを、GOPに分けるようにする。このようにすることで、上位レイヤと下位レイヤとのGOVが1対1に対応することになり、迅速なランダムアクセスが妨げられることを防止することができる。

【0249】即ち、例えば、図39に示した場合と同様に、上位レイヤおよび下位レイヤが構成される場合において、下位レイヤの $P_{1,0}$ が、最初に表示されるVOPとなるように、下位レイヤのGOVを構成したときには、図41に示すように、その $P_{1,0}$ の表示時刻と同時刻に表示される上位レイヤの $P_{2,0}$ が最初に表示されるVOPとなるように、上位レイヤのGOVを構成する。また、例えば、下位レイヤの $P_{1,0}$ が、最初に表示されるVOPとなるように、下位レイヤのGOVを構成したときにも、図41に示すように、その $P_{1,0}$ の表示時刻と同時刻に表示される上位レイヤの $P_{2,0}$ が最初に表示されるものとなるように、上位レイヤのGOVを構成する。

【0250】さらに、例えば、図40に示した場合と同様に、上位レイヤおよび下位レイヤが構成される場合において、下位レイヤの b_1 が、最初に表示されるVOPとなるように、下位レイヤのGOVを構成したときには、図42に示すように、その b_1 の表示時刻の直後に

50

表示される上位レイヤの e_1 が最初に表示されるVOPとなるように、上位レイヤのGOVを構成する。また、例えば、下位レイヤの b_1 が、最初に表示されるVOPとなるように、下位レイヤのGOVを構成したときには、図42に示すように、その b_1 の表示時刻の直後に表示される上位レイヤの e_1 が最初に表示されるVOPとなるように、上位レイヤのGOVを構成する。

【0251】なお、GOVについては、GOVヘッダが各GOV毎に符号化されるが、そのエンコード結果は、GOVの中で最初に符号化／復号されるVOPの直前に配置される。そして、GOVヘッダのtime_codeとしては、GOVの中で最初に表示されるVOPの表示時刻が秒精度でエンコードされる。

【0252】従って、図41において、例えば、 $P_{1,0}$ が最初に表示される下位レイヤのGOVのGOVヘッダに配置されるtime_codeは、その $P_{1,0}$ の表示時刻の秒精度となり、 $P_{2,0}$ が最初に表示される上位レイヤのGOVのGOVヘッダに配置されるtime_codeは、 $P_{2,0}$ の表示時刻の秒精度となる。そして、図41では、 $P_{1,0}$ と $P_{2,0}$ の表示時刻は同一であり、よって、 $P_{1,0}$ が最初に表示される下位レイヤのGOVのGOVヘッダに配置されるtime_codeと、 $P_{2,0}$ が最初に表示される上位レイヤのGOVのGOVヘッダに配置されるtime_codeも同一の値となる。

【0253】また、図42において、例えば、 b_1 が最初に表示される下位レイヤのGOVのGOVヘッダに配置されるtime_codeは、 b_1 の表示時刻の秒精度となり、 e_1 が最初に表示される上位レイヤのGOVのGOVヘッダに配置されるtime_codeは、 e_1 の表示時刻の秒精度となる。

【0254】次に、下位レイヤのGOVを構成する処理（GOVヘッダの挿入処理）は、図9に示した下位レイヤ符号化部25のVLC器36（以下、適宜、下位VLC器36という）において、また、下位レイヤのGOVの最初に表示されるVOPの表示時刻と同時刻またはその直後に表示される上位レイヤのVOPが、上位レイヤのGOVの最初に表示されるものとなるように、上位レイヤのGOPを構成する処理は、図10に示した上位レイヤ符号化部23のVLC器36（以下、適宜、上位VLC器36という）において、それぞれ行われるようになされている。

【0255】そこで、図43または図44のフローチャートを参照して、下位VLC器36または上位VLC器36が行うGOVの構成処理について説明する。

【0256】まず、図43のフローチャートを参照して、下位VLC器36によるGOVの構成処理について説明する。

【0257】下位VLC器36では、まず最初に、ステップS1において、いま処理の対象となっている下位レイヤのVOPのデータが、GOVの中で最初にエンコードされたものであるかどうかを判定され、最初にエンコ

ードされたものでない判定された場合、次の下位レイヤのVOPのデータが供給されるのを待って、ステップS1に戻る。また、ステップS1において、処理の対象となっている下位レイヤのVOPのデータが、GOVの中で最初にエンコードされたものであると判定された場合、ステップS2に進み、その最初にエンコードされたVOPのデータの前に、GOVヘッダが挿入され、ステップS3に進む。ステップS3では、下位レイヤにGOVヘッダを挿入した旨のGOVヘッダ挿入信号が、その下位レイヤのGOVの中で最初に表示されるVOPの表示時刻とともに、上位VLC器36に出力され、次の下位レイヤのVOPのデータが供給されるのを待って、ステップS1に戻る。

【0258】次に、図44のフローチャートを参照して、上位VLC器によるGOVの構成処理について説明する。

【0259】上位VLC器36では、まず最初に、ステップS11において、下位VLC器36から、GOVヘッダ挿入信号が送信されてきたかどうか判定され、送信されてきていないと判定された場合、ステップS11に戻る。また、ステップS11において、GOVヘッダ挿入信号が送信されてきたと判定された場合、ステップS12に進み、そのGOVヘッダ挿入信号とともに送信されてくる、下位レイヤのGOVの中で最初に表示されるVOPの表示時刻が認識される。さらに、ステップS12において、その表示時刻と同時刻またはその直後に表示される上位レイヤのVOPが、上位レイヤのGOPの最初に表示されるものとなるように、GOVヘッダが、符号化ビットストリームに挿入され、ステップS11に戻る。

【0260】なお、図13のデコーダでは、ランダムアクセスが指令された場合、符号化ビットストリームの中のGOVヘッダに配置されたtime_code、並びにVOPヘッダに配置されたmodulo_time_baseおよびVOP_time_incrementを用いて、上述したように、ランダムアクセスすべきVOPの表示時刻が求められ、その表示時刻に基づき、ランダムアクセスが行われる。

【0261】次に、MPEG4では、下位レイヤおよび上位レイヤを構成するVOPのピクチャタイプのシーケンス（ピクチャストラクチャ）については、特に規定されていない。

【0262】そこで、図1のエンコーダでは、例えば、符号化対象の画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを表現するための、例えば、下位レイヤと上位レイヤなどの2以上の階層に階層化した場合には、上位レイヤのVOPを、その表示順序と同一の順序で符号化するようになされている。

【0263】即ち、図45は、空間スケーラブル符号化における上位レイヤと下位レイヤのピクチャストラクチャを示している。なお、図45においては、下位レイヤ

の、表示順で#i番目のI-VOP、P-VOP、またはB-VOPを、それぞれ、 I_{i-1} 、 P_{i-1} 、または B_{i-1} とそれぞれ表し、上位レイヤの、表示順で#i番目のP-VOPまたはB-VOPを、それぞれ、 P_i 、または B_i とそれぞれ表してある。

【0264】図45では、I-VOPである下位レイヤの I_{i-1} が、最初に符号化/復号される。その後、この下位レイヤの I_{i-1} を参照画像として用い、下位レイヤの最初に表示される B_{i-1} が符号化/復号される。空間スケーラブル符号化の上位レイヤのVOPは、基本的に、同時刻に表示される下位レイヤのVOPが参照画像として用いられるため、下位レイヤの B_{i-1} の符号化/復号後は、それを参照画像として用いる上位レイヤの B_{i-1} が符号化/復号される。続いて、下位レイヤの2番目に表示される B_i が、既に符号化/復号されている下位レイヤの I_{i-1} を参照画像として符号化/復号され、その同時刻に表示される上位レイヤである B_i が、下位レイヤの B_{i-1} と、上位レイヤの B_{i-1} を参照画像として符号化/復号される。

【0265】そして、上位レイヤの P_{i-1} が、下位レイヤの I_{i-1} を参照画像として用いて符号化/復号される。その後、下位レイヤ B_{i-1} の符号化/復号のための参照画像となる下位レイヤの P_{i-1} が符号化/復号され、その下位レイヤの P_{i-1} と、同じく下位レイヤの I_{i-1} を参照画像として用いて、下位レイヤの B_i が符号化/復号される。そして、その下位レイヤの B_i と、上位レイヤの P_{i-1} を参照画像として用いて、上位レイヤの B_i が符号化/復号される。

【0266】その後、下位レイヤの B_{i-1} が、下位レイヤの I_{i-1} および P_{i-1} を参照画像として符号化/復号され、その下位レイヤの B_{i-1} と、上位レイヤの B_{i-1} を参照画像として、上位レイヤの B_{i-1} が符号化/復号される。そして、下位レイヤの P_{i-1} は、既に符号化/復号されているので、その下位レイヤの P_{i-1} と、上位レイヤの B_{i-1} を参照画像として、上位レイヤの B_i が符号化/復号される。

【0267】その後、下位レイヤの B_i の参照画像として用いられる下位レイヤの P_{i-1} が符号化/復号され、その下位レイヤの P_{i-1} と、同じく下位レイヤの P_{i-1} を参照画像として、下位レイヤの B_i が符号化/復号される。続いて、その下位レイヤの B_i と、上位レイヤの B_{i-1} を参照画像として、上位レイヤの B_i が符号化/復号され、既に符号化/復号されている下位レイヤの P_{i-1} および P_{i-1} を参照画像として、下位レイヤの B_i が符号化/復号される。そして、その下位レイヤの B_i と、上位レイヤの B_{i-1} を参照画像として、上位レイヤの B_i が符号化/復号され、その上位レイヤの B_{i-1} と、下位レイヤの P_{i-1} を参照画像として、上位レイヤの B_i が符号化/復号される。

【0268】従って、下位レイヤに注目すれば、

B_{1,1}, B_{1,2}, I_{1,1}, B_{1,3}, B_{1,4}, P_{1,1}, B_{1,5}, B_{1,6}, P_{1,2}の順で表示されるVOPが、I_{1,1}, B_{1,1}, B_{1,2}, P_{1,1}, B_{1,3}, B_{1,4}, P_{1,2}, B_{1,5}, B_{1,6}の順で符号化/復号されていく。

【0269】一方、上位レイヤに注目した場合には、B_{1,1}, B_{1,2}, P_{1,1}, B_{1,3}, B_{1,4}, B_{1,5}, B_{1,6}, B_{1,7}, B_{1,8}, B_{1,9}の順で表示されるVOPが、やはり、B_{1,1}, B_{1,2}, P_{1,1}, B_{1,3}, B_{1,4}, B_{1,5}, B_{1,6}, B_{1,7}, B_{1,8}の順で符号化/復号されていく。即ち、上位レイヤのVOPは、その表示順序と同一の順序で符号化/復号される。

【0270】このように、上位レイヤのVOPを、その表示順序と同一の順序で符号化/復号するようにした場合には、表示順序と異なる順序で符号化/復号を行う場合に比較して、処理の容易化を図ることができる。

【0271】ここで、空間スケーラブル符号化を行う場合でも、図1のエンコーダ/図13のデコーダには、図45に示したピクチャストラクチャと異なるピクチャストラクチャを対象に、符号化/復号を行わせることが可能である。即ち、図1のエンコーダ/図13のデコーダに符号化/復号を行わせるピクチャストラクチャは、図45に示したものに限定されるものではない。同様に、参照関係や符号化/復号順序も、図45に示したものに限定されるものではない。

【0272】なお、図45に示した参照関係は、MPEG4のref_select_codeで規定される参照関係を満たすものである。また、空間スケーラブル符号化を行う場合に、上位レイヤのVOPを、その表示順序と同一の順序で符号化/復号する手法は、GOVの有無にかかわらず適用可能である。

【0273】ところで、図45に示したピクチャストラクチャにおいては、上位レイヤの3番目に表示されるP_{1,1}の後に表示されるVOPは、すべてB-VOPになっている。

【0274】一方、VOPの表示時刻は、VOPヘッダのmodulo_time_base, VOP_time_incrementに基づき、GOVヘッダのtime_code。または直前に表示されるI/P-VOPのmodulo_time_baseによって示される同期点を基準として求められる(但し、直前に表示されたI/P-VOPのmodulo_time_baseによって示される同期点、即ち、直前に表示されたI/P-VOPの秒精度の表示時刻も、元をたどれば、GOVヘッダのtime_codeによって示される同期点を基準として求められるから、すべてのVOPの表示時刻は、結局は、GOVヘッダのtime_codeによって示される同期点を基準として求められるといえることができる)。

【0275】従って、図45の下位レイヤについては、B_{1,1}, B_{1,2}, I_{1,1}の表示時刻は、GOVヘッダのtime_codeを、B_{1,3}, B_{1,4}, P_{1,1}の表示時刻は、その直前に表示されるI_{1,1}の秒精度の表示時刻を、B_{1,5}, B_{1,6}, P_{1,2}の表示時刻は、その直前に表示されるP_{1,1}の秒精度の表示時刻を、それぞれ用いて計算される。

【0276】一方、図45の上位レイヤについては、B_{1,1}, B_{1,2}, P_{1,1}の表示時刻は、GOVヘッダのtime_codeを用いて計算される。そして、P_{1,1}の後に表示されるVOPは、上述したように、すべてB-VOPであるから、そのB-VOPの表示時刻は、直前に表示されるP_{1,1}の秒精度の表示時刻を用いて計算される。その結果、P_{1,1}の後に表示されるB-VOPのmodulo_time_baseのビット数が増大し、符号化効率が劣化することになる。

【0277】即ち、注目している注目VOPのmodulo_time_baseは、最下位ビットに0を配置し、その上位ビットに、直前に表示されるI/P-VOPの表示時刻の秒精度と、注目VOPの表示時刻の秒精度との差分と同一の数の1を配置して構成される。従って、modulo_time_baseは、I/P-VOPがあると、その後に、いわば"0"にリセットされ、これにより、そのビット数の増大の防止が図られている。

【0278】しかしながら、図45に示した上位レイヤについては、P_{1,1}の後に表示されるVOPは、すべてB-VOPであるから、modulo_time_baseのリセットが行われず、そのビット数が増大することになる。

【0279】具体的には、図46に示すように、下位レイヤについては、I_{1,1}, P_{1,1}, P_{1,2}それぞれの後において、modulo_time_baseは、"0"にリセットされる。そして、I_{1,1}の後のB_{1,1}, P_{1,1}の後のB_{1,2}, またはP_{1,1}の後のB_{1,3}については、I_{1,1}, P_{1,1}, P_{1,2}から、B_{1,1}, B_{1,2}, B_{1,3}までのそれぞれの表示間隔の秒精度に対応した数の"1"が"0"にリセットされたmodulo_time_baseの上位ビットに付される(図46では、I_{1,1}, P_{1,1}, P_{1,2}から、B_{1,1}, B_{1,2}, B_{1,3}までのそれぞれの表示間隔は、いずれも1秒未満であるため、それぞれのmodulo_time_baseは、"0"の上位ビットとして0個の"1"が付された"0"となっている)。

【0280】これに対して、上位レイヤについては、P_{1,1}の後において、modulo_time_baseが"0"にリセットされた後は、P_{1,1}の秒精度の表示時刻を基準とした秒精度の表示時刻が、1秒増加することに、VOPに付されるmodulo_time_baseも、1ビットずつ増加していく。その結果、図46に示すように、P_{1,1}の後においてmodulo_time_baseが"0"にリセットされた後、その後に表示されるB_{1,1}, B_{1,2}, B_{1,3}, B_{1,4}, B_{1,5}, B_{1,6}, B_{1,7}のmodulo_time_baseは、それぞれ"0", "10", "10", "110", "110", "1110", "1110"となり、そのビット数が増大していく。

【0281】そこで、例えば、図45に示したように、空間スケーラブル符号化において、上位レイヤの符号化/復号が表示順序と同一順序で行われる場合には、上位レイヤのB-VOPの後においても、modulo_time_baseを"0"にリセットし、そのB-VOPと、次に表示されるVOPとの秒精度の表示時刻の差分と同一の数の"1"を、"0"の上位ビットとして付加して、次に表示されるVOP

(29)

特開 2000-13790

55

のmodulo_time_baseを構成するようにする。この場合、上位レイヤのVOPの表示時刻は、そのVOPの直前に表示されるVOP (I, P, BのいずれのVOPであっても良い) の秒精度の表示時刻を基準にして計算されることになる。なお、注目VOPの前に表示されるVOPが存在しない場合、即ち、注目VOPが、GOVにおいて、最初に表示されるVOPである場合には、その直前に表示されるVOPは存在しないため、この最初に表示されるVOPについては、例えば、time_codeを基準として、modulo_time_baseの構成、および表示時刻の計算を行うこととする。

【0282】図47は、modulo_time_baseの構成方法およびVOPの表示時刻の計算方法を、上述のように変更した場合における、図45に示したピクチャストラクチャのmodulo_time_baseを示している。

【0283】この場合、上位レイヤの最初に表示されるB₁については、その表示時刻の秒精度 (図47では、0h:12m:34s) と、time_code (図47では、0h:12m:34s) との差分が0となるから、そのmodulo_time_baseは、"0"に、0個の"1"を付加した"0"となる。B₂についても、その表示時刻の秒精度 (図47では、0h:12m:34s) と、その直前に表示されるB₁の表示時刻の秒精度 (図47では、0h:12m:34s) との差分が0となるから、そのmodulo_time_baseは、"0"に、0個の"1"を付加した"0"となる。P₁については、その表示時刻の秒精度 (図47では、0h:12m:35s) と、その直前に表示されるB₂の表示時刻の秒精度 (図47では、0h:12m:34s) との差分が1となるから、そのmodulo_time_baseは、"0"に、1個の"1"を付加した"10"となる。B₃については、その表示時刻の秒精度 (図47では、0h:12m:35s) と、その直前に表示されるP₁の表示時刻の秒精度 (図47では、0h:12m:35s) との差分が0となるから、そのmodulo_time_baseは、"0"に、0個の"1"を付加した"0"となる。即ち、B₃の表示時刻の秒精度と、その直前に表示されるP₁の表示時刻の秒精度との差分が1秒に満たないため、いわば、B₃において、modulo_time_baseはリセットされる。

【0284】B₄については、その表示時刻の秒精度 (図47では、0h:12m:36s) と、その直前に表示されるB₃の表示時刻の秒精度 (図47では、0h:12m:35s) との差分が1となるから、そのmodulo_time_baseは、"0"に、1個の"1"を付加した"0"となる。B₅については、その表示時刻の秒精度 (図47では、0h:12m:36s) と、その直前に表示されるB₄の表示時刻の秒精度 (図47では、0h:12m:36s) との差分が0となるから、そのmodulo_time_baseは、B₄における場合と同様に、"0"にリセットされる。

【0285】以下、同様にして、B₆, B₇, B₈, B₉のmodulo_time_baseは、図47に示すように、それぞれ、"10", "0", "10", "0"となる。

【0286】次に、modulo_time_baseを設定する処理

56

(modulo_time_base設定処理)は、下位レイヤまたは上位レイヤそれぞれについて、図9に示した下位レイヤ符号化部25のVLC器36 (下位VLC器36)、または図10に示した上位レイヤ符号化部23のVLC器36 (上位VLC器36)において、それぞれ行われるようになされている。

【0287】そこで、図48のフローチャートを参照して、上位VLC器36が行うmodulo_time_base設定処理について説明する。

【0288】上位VLC器36では、いま処理の対象となっているVOPを注目VOPとして、ステップS21において、その注目VOPと、直前に表示されるVOPとの秒精度の表示時刻の差分Dが求められ、ステップS22に進み、Dが0に等しいかどうか判定される。ステップS22において、Dが0に等しいと判定された場合、ステップS23に進み、modulo_time_baseが"0"にリセットされ、ステップS28に進む。ステップS28では、"0"にリセットされたmodulo_time_baseが、注目VOPに設定され、次の処理対象となるVOPが供給されるのを待って、ステップS21に戻る。

【0289】一方、ステップS22において、Dが0に等しくない判定された場合、ステップS24に進み、modulo_time_baseが"0"にリセットされる。そして、ステップS25において、modulo_time_baseの最上位ビットとして、"1"が追加され、ステップS26に進む。ステップS26では、Dが1だけデクリメントされ、ステップS27に進み、Dが0に等しいかどうか判定される。ステップS27において、Dが0に等しくない判定された場合、ステップS25に戻り、以下、ステップS27でDが0に等しいと判定されるまで、ステップS25乃至S27の処理が繰り返される。

【0290】そして、ステップS27において、Dが0に等しいと判定された場合、即ち、注目VOPと、直前に表示されるVOPとの秒精度の表示時刻の差分と同一の数の"1"が、"0"に付加されたmodulo_time_baseが構成された場合、ステップS28に進み、そのmodulo_time_baseが、注目VOPに設定され、次の処理対象となるVOPが供給されるのを待って、ステップS21に戻る。

【0291】なお、下位VLC器36では、図48のステップS21において、注目VOPと、直前に表示されるVOP (I/P/B-VOP) との秒精度の表示時刻の差分ではなく、注目VOPと、直前に表示されるI/P-VOPとの秒精度の表示時刻の差分Dが求められることを除けば、上位VLC器36と同様のmodulo_time_base設定処理が行われる。

【0292】また、図13のデコーダでは、下位レイヤのVOPの表示時刻は、その直前に表示されるI/P-VOPの秒精度の表示時刻 (またはtime_code) を基準として求められるが、上位レイヤのVOPの表示時刻は、その直前に表示されるVOP (I, P, Bを問わない) の秒精度の表

(30)

特開 2000-13790

57

示時刻（またはtime_code）を基準として求められる。

【0293】ところで、図45に示した場合においては、上位レイヤの B_{i-1} の符号化／復号は、同時刻に表示される下位レイヤの B_{i-1} のみを参照画像として行われる。

【0294】上位レイヤの B_{i-1} は、B-VOPであるが、B-VOPは、通常、I/P-VOPよりも発生ビット量が少なくなるように符号化されるため、 B_{i-1} の復号画像の画質は劣化する可能性が高い。さらに、 B_{i-1} の復号画像の画質が劣化する場合には、その B_{i-1} を参照画像として用いる B_{i-1} の復号画像の画質も劣化する可能性が高い。即ち、図45に示した場合では、上位レイヤにおいて、P-VOPである P_{i-1} が表示される前に表示されるB-VOPの画質は劣化する可能性が高い。

【0295】そこで、そのような画質の劣化を防止（低減）するために、例えば、図45に示した、空間スケラビリティを実現するピクチャストラクチャについては、その符号化／復号順序と参照関係を、例えば、図49に示すようにすることができる。

【0296】即ち、図49では、下位レイヤのI-VOPを符号化／復号した場合には、上位レイヤにおいて、下位レイヤのI-VOPの表示時刻と同時刻に表示されるP-VOPが、下位レイヤのI-VOPの直後に符号化／復号される。

【0297】従って、図49に示した場合においては、まず最初に、下位レイヤの最初に表示されるI-VOPである I_{i-1} が符号化／復号される。そして、下位レイヤの I_{i-1} と同時刻に表示される上位レイヤのP-VOPである P_{i-1} が、直前に符号化／復号された I_{i-1} を参照画像として用いて符号化／復号される。

【0298】その後、さらに、下位レイヤの I_{i-1} を参照画像として用いて、下位レイヤの B_{i-1} が符号化／復号される。次に、上位レイヤの B_{i-1} が符号化／復号されるが、この上位レイヤの B_{i-1} の符号化／復号には、下位レイヤの B_{i-1} だけでなく、上位レイヤの P_{i-1} も、参照画像として用いられる。これは、MPEG4では、空間スケラブル符号化におけるB-VOPについては、前方予測（forward prediction）につき、「most recent decoded enhancement VOP of the same layer」を参照画像として用いることが定められているためであり、図49における場合、「most recent decoded enhancement VOP of the same layer」は、 P_{i-1} であるからである。

【0299】上位レイヤの B_{i-1} の符号化／復号後は、下位レイヤの B_{i-1} が、同じく下位レイヤの I_{i-1} を参照画像として符号化／復号され、その後、上位レイヤの B_{i-1} および下位レイヤの B_{i-1} を参照画像として用いて、上位レイヤの P_{i-1} が符号化／復号される。続いて、下位レイヤの P_{i-1} が符号化／復号され、以下、図45における場合と同様に、下位レイヤの B_{i-1} 、上位レイヤの B_{i-1} 、下位レイヤの B_{i-1} 、上位レイヤの B_{i-1} 、下位レイヤの B_{i-1} 、上位レイヤの B_{i-1} 、下位レイヤ

58

の B_{i-1} 、上位レイヤの B_{i-1} 、 B_{i-1} の順で符号化／復号される。

【0300】以上のように、下位レイヤのI-VOPを符号化／復号した場合には、上位レイヤにおいて、下位レイヤのI-VOPの表示時刻と同時刻に表示されるP-VOPを、下位レイヤのI-VOPの直後に符号化／復号するようにすることで、上位レイヤにおいては、そのP-VOPを参照画像として、最初に表示されるB-VOPが符号化／復号されることになるので（図49では、 P_{i-1} を参照画像として、 B_{i-1} が符号化／復号されることになるので）、上位レイヤにおいて、P-VOPの前に表示されるB-VOPの画質の劣化を防止することが可能となる。

【0301】ここで、図1のエンコーダ／図13のデコーダには、図49に示した参照関係や符号化／復号順序以外の参照関係や符号化／復号順序に基づいて、符号化／復号を行わせることが可能である。即ち、参照関係や符号化／復号順序は、図49に示したものに限定されるものではない。

【0302】なお、図49に示した参照関係は、MPEG4のref_select_codeで規定される参照関係を満たすものである。また、空間スケラブル符号化を行う場合に、図49に示した順序で符号化／復号を行う手法は、GOVの有無にかかわらず適用可能である。

【0303】さらに、図49に示した参照関係にしたがって符号化／復号を行う場合には、その符号化／復号順序は、上述した場合と異なる順序とすることが可能である。即ち、例えば、下位レイヤのすべてのVOPを符号化／復号した後に、上位レイヤのVOPの符号化／復号を行うようにすることなどが可能である。但し、図49では、上位レイヤのVOPの符号化／復号のための参照画像として、下位レイヤのVOPを用いているため、この場合、上位レイヤのVOPの符号化／復号のための参照画像として用いる下位レイヤのVOPを、その上位レイヤのVOPの符号化／復号が終了するまで、長時間、保持することが必要となることがある。そのような長時間の保持を避けるためには、図49で説明したような符号化／復号順序を採用するのが望ましい。なお、このことは、図45に示した場合についても、同様である。

【0304】次に、本実施の形態では、迅速で、効率的なランダムアクセスが妨げられることを防止するために、図39乃至図42で説明したように、下位レイヤのGOVの最初に表示されるVOPの表示時刻と同時刻またはその直後に表示される上位レイヤのVOPが、上位レイヤのGOVの最初に表示されるものとなるように、上位レイヤを、GOVに分けるようにした。

【0305】しかしながら、GOVヘッダのtime_codeには、GOVの中で最初に表示されるVOPの表示時刻を秒精度で設定するようにしたため、下位レイヤのGOVの最初に表示されるVOPの表示時刻の直後に表示される上位レイヤのVOPが、上位レイヤのGOVの最初に表示さ

れるものとなるように、上位レイヤのGOVを構成した場合、下位レイヤのGOVのtime_codeに設定される値(秒精度先頭表示時刻)と、その下位レイヤのGOVに対応する上位レイヤのGOVのtime_codeに設定される値とが一致しないことがある。

【0306】即ち、図50は、時間的スケラブル符号化を行う場合の、下位レイヤおよび上位レイヤの構成例を示している。

【0307】いま、図50において、下位レイヤのB₁を、最初に表示されるVOPとして、GOVを構成した場合、上位レイヤについては、下位レイヤのGOVの最初に表示されるB₁の表示時刻の直後に表示される上位レイヤのVOPであるB₂が最初に表示されるVOPとなるように、上位レイヤのGOPが構成される。この場合、下位レイヤで構成されるGOVのGOVヘッダのtime_codeには、そのGOVの中で最初に表示されるB₁の表示時刻が秒精度で設定される。同様に、上位レイヤで構成されるGOVのGOVヘッダのtime_codeには、そのGOVの中で最初に表示されるB₂の表示時刻が秒精度で設定される。

【0308】即ち、図50では、B₁の表示時刻は、01h:12m:33sから01h:12m:34sの間の時刻であり、B₂の表示時刻は、01h:12m:34sから01h:12m:35sの間の時刻であるから、下位レイヤで構成されるGOVのGOVヘッダのtime_codeには、01h:12m:33sが設定され、上位レイヤで構成されるGOVのGOVヘッダのtime_codeには、01h:12m:34sが設定される。従って、下位レイヤのGOVのtime_codeに設定される値と、その下位レイヤのGOVに対応する上位レイヤのGOVのtime_codeに設定される値とは一致しない。

【0309】そして、このように、下位レイヤのGOVのtime_codeに設定される値と、その下位レイヤのGOVに対応する上位レイヤのGOVのtime_codeに設定される値とが一致しない場合には、符号化ビットストリームにランダムアクセスする際に、下位レイヤと上位レイヤとにおいて、異なる秒精度の時刻(time_codeによって示される同期点)を基準に、VOPの表示時刻を計算しなければならないことになる。

【0310】そこで、上位レイヤのGOVのtime_codeには、その上位レイヤに対応する下位レイヤのGOVのtime_codeに設定される値と同一の値を設定するようにすることができる。

【0311】即ち、図51は、図50における場合と同様の時間的スケラブル符号化を行う場合の、下位レイヤおよび上位レイヤの構成例を示している。

【0312】いま、図51において、図50における場合と同様に、下位レイヤのB₁を、最初に表示されるVOPとして、GOVを構成した場合、上位レイヤについては、下位レイヤのGOVの最初に表示されるB₁の表示時刻の直後に表示される上位レイヤのVOPであるB₂が最初に表示されるVOPとなるように、上位レイヤのG

OPが構成される。

【0313】そして、この場合、下位レイヤで構成されるGOVのGOVヘッダのtime_codeには、図50における場合と同様に、そのGOVの中で最初に表示されるB₁の表示時刻を秒精度で設定する。従って、下位レイヤで構成されるGOVのGOVヘッダのtime_codeには、01h:12m:33sが設定される。

【0314】一方、上位レイヤで構成されるGOVのGOVヘッダのtime_codeには、その上位レイヤのGOVに対応する下位レイヤのGOVのtime_codeに設定される値と同一の値を設定する。従って、下位レイヤで構成されるGOVのGOVヘッダのtime_codeにも、01h:12m:33sが設定される。

【0315】この場合、符号化ビットストリームにランダムアクセスする際に、下位レイヤと上位レイヤとにおいて、異なる秒精度の時刻(time_codeによって示される同期点)を基準に、VOPの表示時刻を計算しなければならないことを防止することができる。即ち、下位レイヤと上位レイヤとにおいて、常に、同一の秒精度の時刻(time_codeによって示される同期点)を基準に、VOPの表示時刻を計算することが可能となる。

【0316】なお、上位レイヤで構成されるGOVのGOVヘッダのtime_codeに、その上位レイヤのGOVに対応する下位レイヤのGOVのtime_codeに設定される値と同一の値を設定する場合も、上位レイヤで構成されるGOVのGOVヘッダのtime_codeに、そのGOVの中で最初に表示されるVOPの表示時刻を秒精度で設定する場合と同様に、図1のエンコーダでは、modulo_time_baseおよびVOP_time_incrementを求めることができ、また、図13のデコーダでは、そのmodulo_time_baseおよびVOP_time_increment、並びにtime_codeを用いて、VOPの表示時刻を求めることができる。

【0317】次に、以上説明したエンコーダおよびデコーダは、それ専用のハードウェアによって実現することもできるし、コンピュータに、上述したような処理を行わせるためのプログラムを実行させることによって実現することができる。

【0318】即ち、図52は、図1のエンコーダまたは図13のデコーダとして機能するコンピュータの一実施の形態の構成例を示している。

【0319】ROM(Read Only Memory)201は、例えば、ブートプログラムなどを記憶している。CPU(Central Processing Unit)202は、例えば、HD(HardDisk)206に記憶されたプログラムを、RAM(Read Only Memory)203上に展開して実行することで、各種の処理を行うようになされている。RAM203は、CPU202が実行するプログラムや、CPU202の処理上必要なデータを一時記憶するようになされている。入力部204は、例えば、キーボードやマウスなどでなり、必要なコマンドやデータを入力するときな

どに操作される。出力部 205 は、例えば、ディスプレイなどでなり、CPU 202 の制御にしたがったデータを表示する。HD 206 は、CPU 202 が実行すべきプログラム、さらには、エンコード対象の画像データや、エンコード後のデータ（符号化ビットストリーム）、デコード後の画像データなどを記憶するようになされている。通信 I/F（Interface）207 は、外部との通信を制御することにより、例えば、エンコード対象の画像データを、外部から受信したり、また、エンコード後の符号化ビットストリームを外部に送信したりするようになされている。また、通信 I/F 207 は、外部でエンコードされた符号化ビットストリームを受信したり、また、デコード後の画像データを、外部に送信するようになされている。

【0320】以上のように構成されるコンピュータの CPU 202 に、上述したような処理を行うためのプログラムを実行させることにより、このコンピュータは、図 1 に示したエンコーダや、図 13 に示したデコーダとして機能することになる。

【0321】なお、本発明の適用範囲は、MPEG4 に準拠した範囲に限られるものではない。

【0322】

【発明の効果】請求項 1 に記載の画像符号化装置および請求項 7 に記載の画像符号化方法によれば、画像を構成するオブジェクトのシーケンスが 2 以上の階層に階層化され、オブジェクトのシーケンスの第 1 の階層が、複数のグループに分けて符号化されるとともに、第 1 の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第 2 の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第 2 の階層が、複数のグループに分けて符号化される。従って、効率的かつ迅速なランダムアクセスが可能となる。

【0323】請求項 8 に記載の画像復号装置および請求項 14 に記載の画像復号方法によれば、画像を構成するオブジェクトのシーケンスを 2 以上の階層に階層化し、オブジェクトのシーケンスの第 1 の階層を、複数のグループに分けて符号化するとともに、第 1 の階層のグループの最初に表示されるオブジェクトの表示時刻と同時刻またはその直後に表示される第 2 の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第 2 の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームが復号される。従って、効率的かつ迅速なランダムアクセスが可能となる。

【0324】請求項 15 に記載の提供媒体によれば、画像を構成するオブジェクトのシーケンスを 2 以上の階層に階層化し、オブジェクトのシーケンスの第 1 の階層を、複数のグループに分けて符号化するとともに、第 1 の階層のグループの最初に表示されるオブジェクトの表

示時刻と同時刻またはその直後に表示される第 2 の階層のオブジェクトが、グループの最初に表示されるオブジェクトとなるように、オブジェクトのシーケンスの第 2 の階層を、複数のグループに分けて符号化することにより得られる符号化ビットストリームが提供される。従って、その符号化ビットストリームに対しては、効率的かつ迅速にランダムアクセスすることができる。

【0325】請求項 21 に記載の画像符号化装置および請求項 22 に記載の画像符号化方法によれば、画像を構成するオブジェクトのシーケンスが 2 以上の階層に階層化され、オブジェクトのシーケンスの第 1 または第 2 の階層が、1 以上のグループに分けてそれぞれ符号化される。さらに、第 1 または第 2 の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻がそれぞれ含められ、第 1 または第 2 の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報が付加される。そして、この場合に、第 2 の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第 2 の階層のオブジェクトについての秒精度相対時刻情報がリセットされる。従って、符号化効率の劣化を防止することが可能となる。

【0326】請求項 23 に記載の画像復号装置および請求項 24 に記載の画像復号方法によれば、画像を構成するオブジェクトのシーケンスを 2 以上の階層に階層化し、オブジェクトのシーケンスの第 1 または第 2 の階層を、1 以上のグループに分けてそれぞれ符号化し、第 1 または第 2 の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第 1 または第 2 の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビットストリームであって、第 2 の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第 2 の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものが復号される。従って、符号化効率の劣化を防止された符号化ビットストリームを復号することができる。

【0327】請求項 25 に記載の提供媒体によれば、画像を構成するオブジェクトのシーケンスを 2 以上の階層に階層化し、オブジェクトのシーケンスの第 1 または第 2 の階層を、1 以上のグループに分けてそれぞれ符号化し、第 1 または第 2 の階層のグループに、その最初に表示されるオブジェクトの表示時刻を秒精度で表す秒精度先頭表示時刻をそれぞれ含め、第 1 または第 2 の階層のオブジェクトそれぞれに、秒精度先頭表示時刻を基準とする、そのオブジェクトの表示時刻の秒精度を表す秒精度相対時刻情報を付加することにより得られる符号化ビ

ットストリームであって、第2の階層の、表示順で隣接するオブジェクトどうしの表示時刻の差に基づいて、第2の階層のオブジェクトについての秒精度相対時刻情報がリセットされているものが提供される。従って、符号化効率の劣化の防止された符号化ビットストリームを提供することができる。

【0328】請求項26に記載の画像符号化装置および請求項27に記載の画像符号化方法によれば、画像を構成するオブジェクトのシーケンスが、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化され、下位階層のオブジェクトのシーケンスが符号化されるとともに、上位階層のオブジェクトのシーケンスが、その表示順序と同一の順序で符号化される。従って、符号化処理の容易化を図ることが可能となる。

【0329】請求項28に記載の画像復号装置および請求項29に記載の画像復号方法によれば、画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームが復号される。従って、復号処理の容易化を係ることが可能となる。

【0330】請求項30に記載の提供媒体によれば、画像を構成するオブジェクトのシーケンスを、空間スケーラビリティを実現するための上位階層と下位階層とを含む2以上の階層に階層化し、下位階層のオブジェクトのシーケンスを符号化するとともに、上位階層のオブジェクトのシーケンスを、その表示順序と同一の順序で符号化することにより得られる符号化ビットストリームが提供される。従って、比較的容易に復号可能な符号化ビットストリームを提供することができる。

【図面の簡単な説明】

【図1】本発明を適用したエンコーダの一実施の形態の構成例を示すブロック図である。

【図2】時刻によって、VOの位置、大きさが変化することを説明するための図である。

【図3】図1のVOP符号化部3、乃至3、の構成例を示すブロック図である。

【図4】空間スケーラビリティを説明するための図である。

【図5】空間スケーラビリティを説明するための図である。

【図6】空間スケーラビリティを説明するための図である。

【図7】空間スケーラビリティを説明するための図である。

【図8】VOPのサイズデータおよびオフセットデータの決定方法を説明するための図である。

【図9】図3の下位レイヤ符号化部25の構成例を示すブロック図である。

【図10】図3の上位レイヤ符号化部23の構成例を示すブロック図である。

【図11】空間スケーラビリティを説明するための図である。

【図12】時間スケーラビリティを説明するための図である。

【図13】本発明を適用したデコーダの一実施の形態の構成例を示すブロック図である。

【図14】図13のVOP復号部72、乃至72、の他の構成例を示すブロック図である。

【図15】図14の下位レイヤ復号部95の構成例を示すブロック図である。

【図16】図14の上位レイヤ復号部93の構成例を示すブロック図である。

【図17】MPEG4規格のVideo Verification Model (Version 10.0)における符号化ビットストリームの構成を示す図である。

【図18】VSのシンタックスを示す図である。

【図19】VOのシンタックスを示す図である。

【図20】VOLのシンタックスを示す図である。

【図21】VOLのシンタックスを示す図である。

【図22】VOLのシンタックスを示す図である。

【図23】VOLのシンタックスを示す図である。

【図24】VOLのシンタックスを示す図である。

【図25】VOLのシンタックスを示す図である。

【図26】GOVのシンタックスを示す図である。

【図27】GOVのシンタックスを示す図である。

【図28】VOPのシンタックスを示す図である。

【図29】VOPのシンタックスを示す図である。

【図30】VOPのシンタックスを示す図である。

【図31】VOPのシンタックスを示す図である。

【図32】VOPのシンタックスを示す図である。

【図33】VOPのシンタックスを示す図である。

【図34】VOPのシンタックスを示す図である。

【図35】VOPのシンタックスを示す図である。

【図36】VOPのシンタックスを示す図である。

【図37】time_code, modulo_time_base, VOP_time_incrementの関係を示す図である。

【図38】time_code, modulo_time_base, VOP_time_incrementの関係を示す図である。

【図39】下位レイヤと上位レイヤとで独立にGOVを構成した状態を示す図である。

【図40】下位レイヤと上位レイヤとで独立にGOVを構成した状態を示す図である。

【図41】上位レイヤのGOVを、下位レイヤのGOVと対応付けて構成した状態を示す図である。

【図42】上位レイヤのGOVを、下位レイヤのGOVと対応付けて構成した状態を示す図である。

(34)

特開 2000-13790

65

66

【図 4 3】上位レイヤの GOV を、下位レイヤの GOV と対応付けて構成する場合の、図 9 の VLC 器 3 6 の処理を説明するためのフローチャートである。

【図 4 4】上位レイヤの GOV を、下位レイヤの GOV と対応付けて構成する場合の、図 10 の VLC 器 3 6 の処理を説明するためのフローチャートである。

【図 4 5】空間スケーラブル符号化における上位レイヤと下位レイヤのピクチャストラクチャの例を示す図である。

【図 4 6】 modulo_time_base のビット数が増大する場合 10 を説明するための図である。

【図 4 7】 modulo_time_base のビット数の増大を防止する方法を説明するための図である。

【図 4 8】 modulo_time_base のビット数の増大を防止する処理を説明するための図である。

【図 4 9】画質の劣化を防止するための符号化／復号順序と参照関係を説明するための図である。

【図 5 0】下位レイヤと上位レイヤとの time_code が一致しない状態を示す図である。

【図 5 1】下位レイヤと上位レイヤとの time_code を一致させる方法を説明するための図である。 20

【図 5 2】本発明を適用したエンコーダおよびデコーダの他の実施の形態の構成例を示すブロック図である。

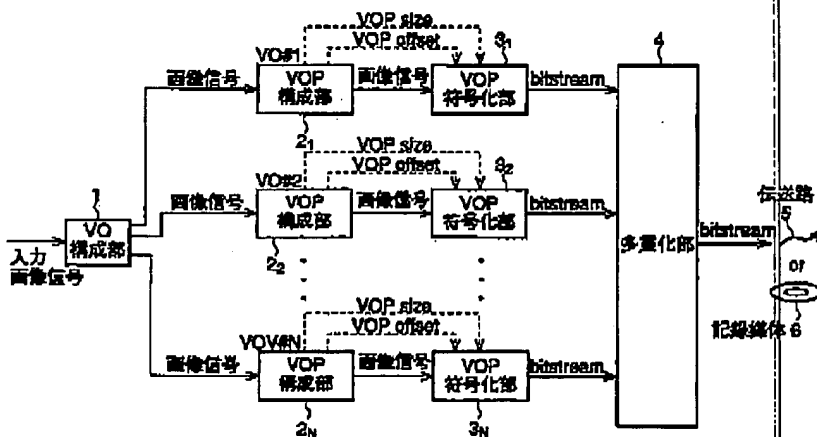
【図 5 3】従来のエンコーダの一例の構成を示すブロック図である。

【図 5 4】従来のデコーダの一例の構成を示すブロック図である。

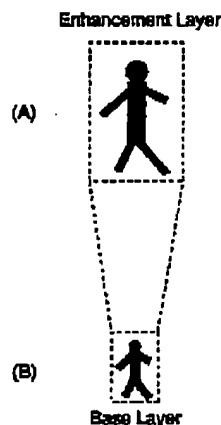
【符号の説明】

1 VOP 構成部, 2, 乃至 2, VOP 構成部, 3, 乃至 3, VOP 符号化部, 4 多重化部, 21 画像階層化部, 23 上位レイヤ符号化部, 24 解像度変換部, 25 下位レイヤ符号化部, 26 多重化部, 31 フレームメモリ, 32 動きベクトル検出器, 33 演算器, 34 DCT 器, 35 量子化器, 36 VLC 器, 38 逆量子化器, 39 IDCT 器, 40 演算器, 41 フレームメモリ, 42 動き補償器, 53 フレームメモリ, 71 逆多重化部, 72, 乃至 72, VOP 復号部, 73 画像再構成部, 91 逆多重化部, 93 上位レイヤ復号部, 94 解像度変換部, 95 下位レイヤ復号部, 102 IVLC 器, 103 逆量子化器, 104 IDCT 器, 105 演算器, 106 フレームメモリ, 107 動き補償器, 112 フレームメモリ, 201 ROM, 202 CPU, 203 RAM, 204 入力部, 205 出力部, 206 HD, 207 通信 I/F

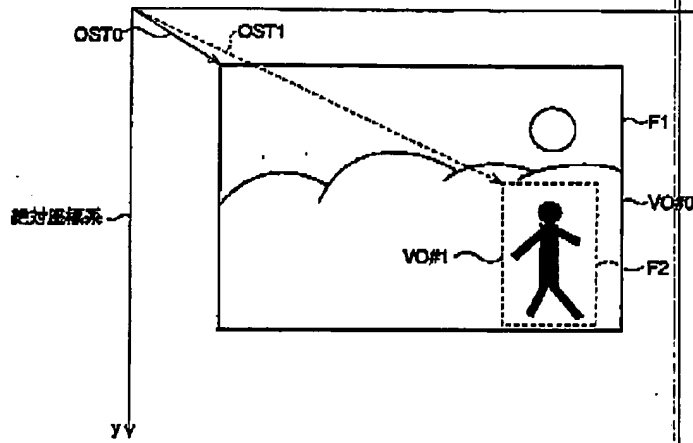
【図 1】



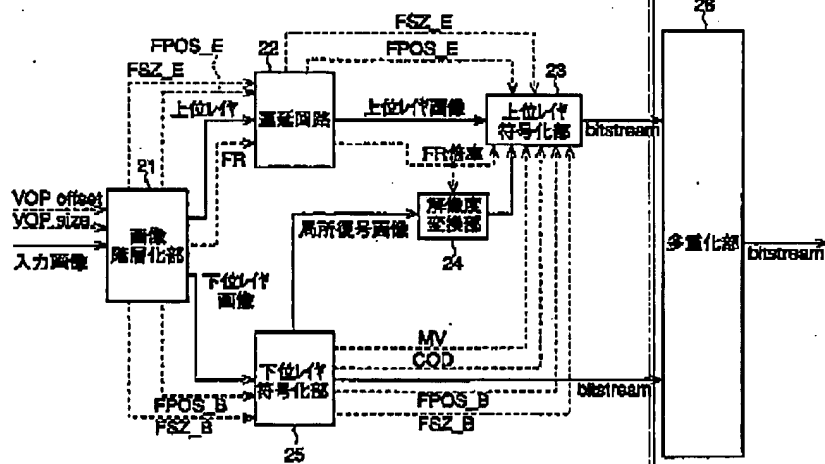
【図 7】



【圖 2】



【图 3】



VOP符号化部 3n

【例 19】

Video Object Class	No. of bits	Min/max
Video Object		
Syntax		
VideoObject () { video_object_start_code video_object_id do { VideoObjectLayer () } while (next_hls_byte_assigned () == video_object_layer_start_code) do { texture_object_layer () } while (next_hls_byte_assigned () == texture_object_layer_start_code) next_start_code () }	go43-27 5	

video_object_start_code

This is a unique code of length 27 bits($8c+3$) that precedes the `video_object_id`.

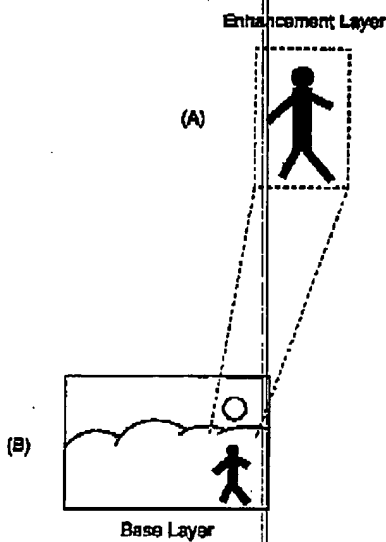
video_object_id

This is a 5-bit code which identifies a video object in a scene being processed.

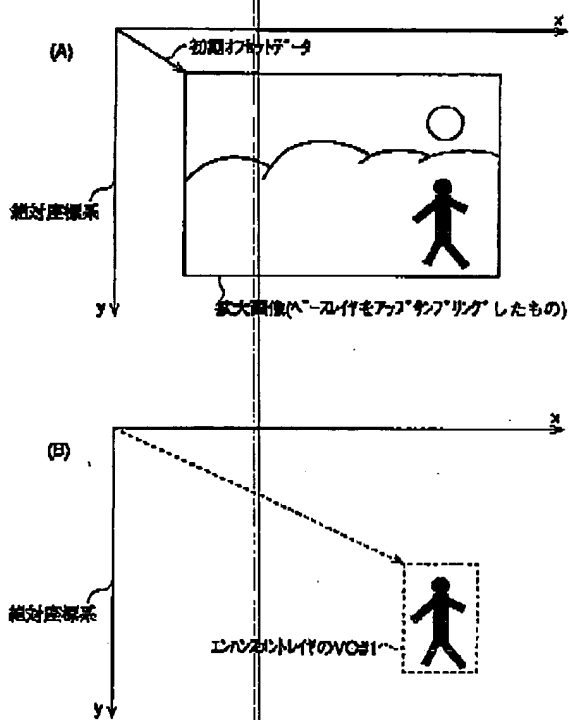
特開2000-13790

(36)

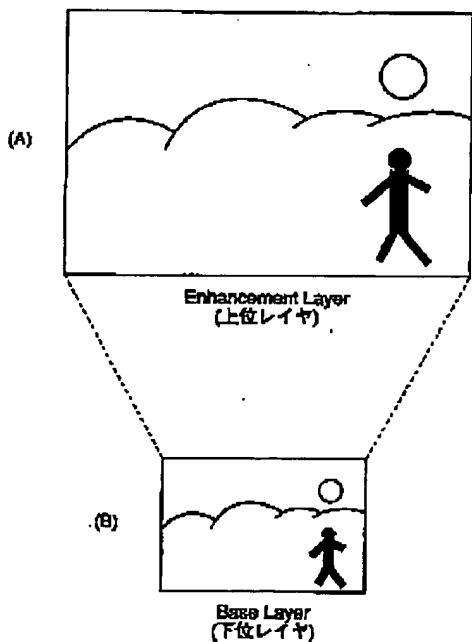
【図5】



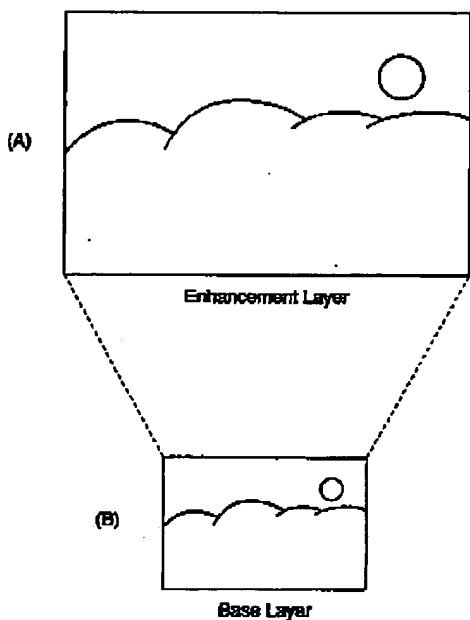
【図8】



【図4】



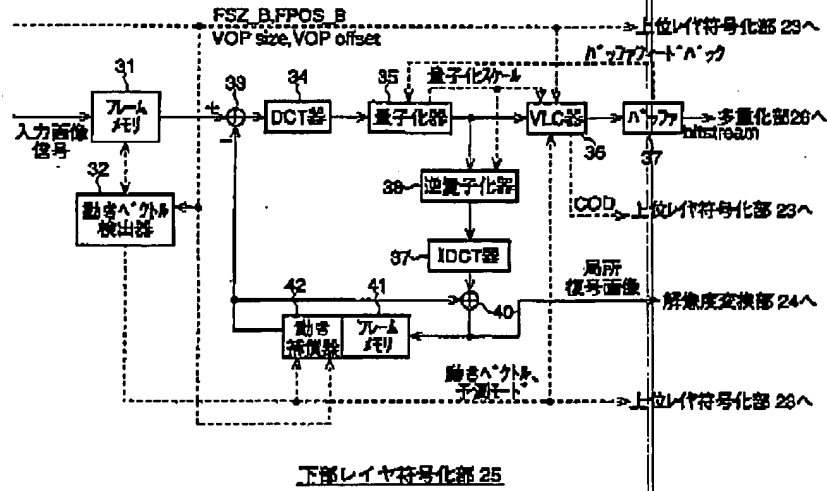
【図6】



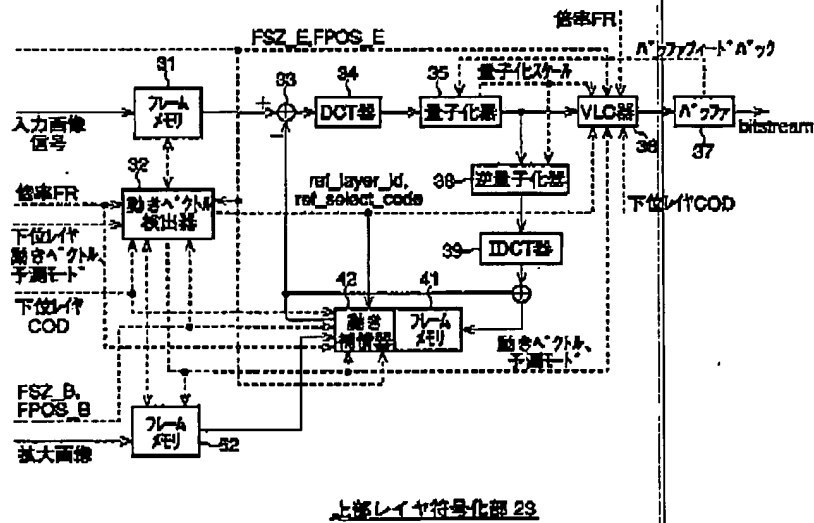
(37)

特開 2000-13790

【図 9】



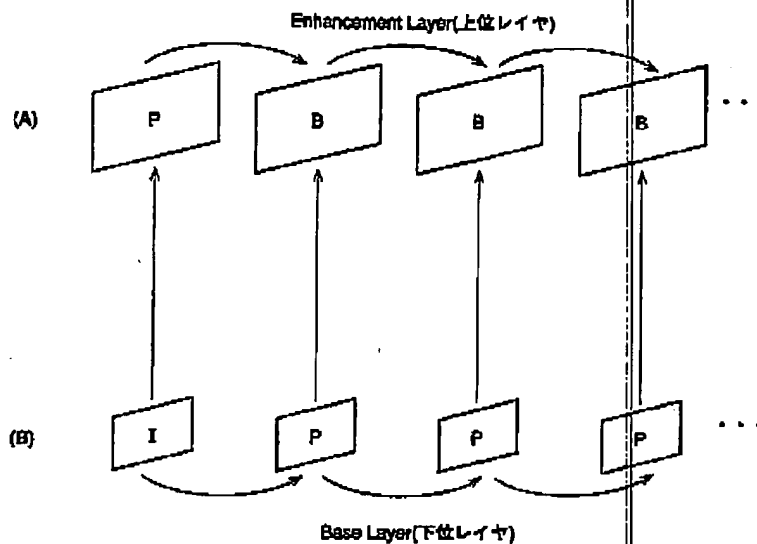
【図 10】



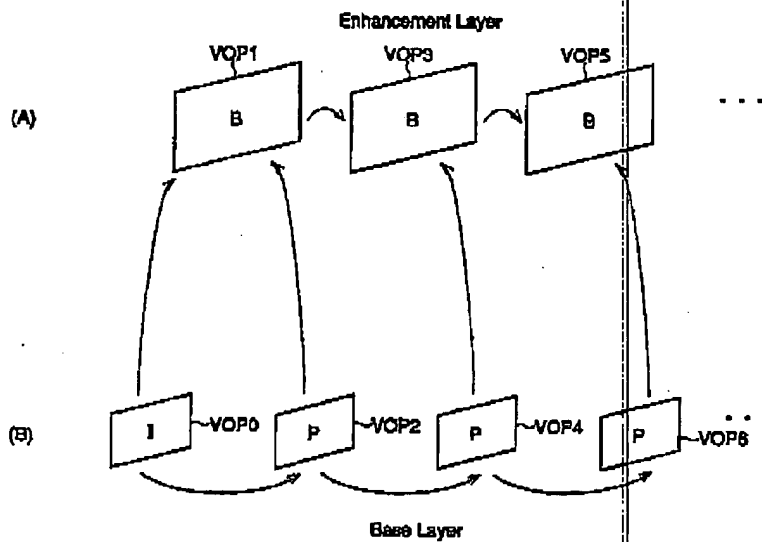
特開2000-13790

(38)

【図11】



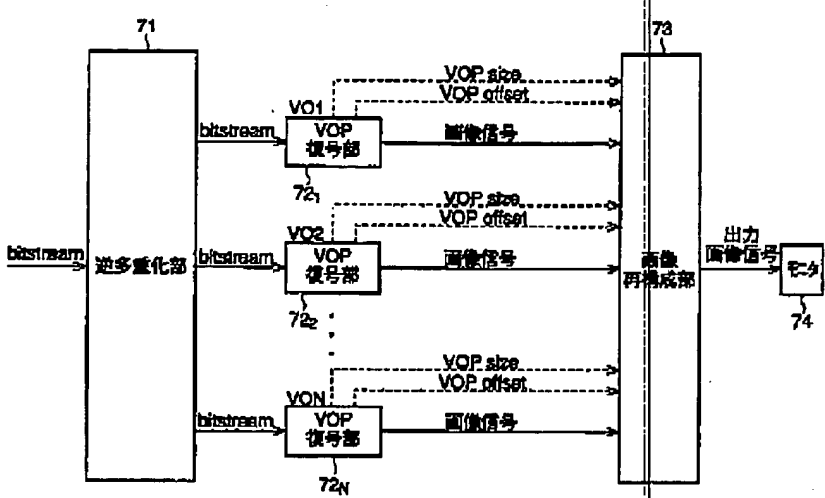
【図12】



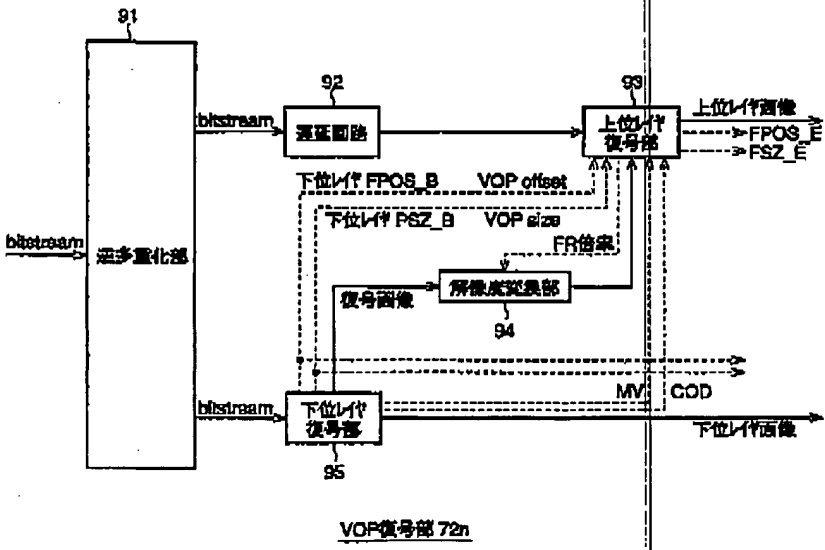
特 許 2 0 0 0 - 1 3 7 9 0

(3 9)

【 図 1 3 】



【 図 1 4 】

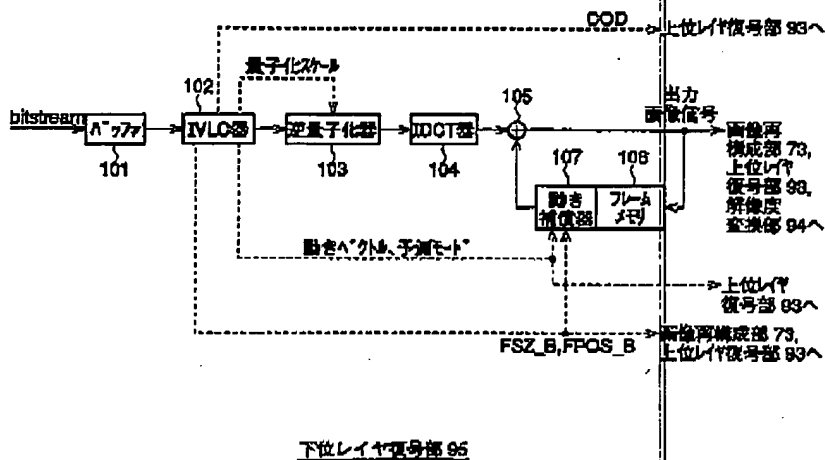


VOP復号部 72n

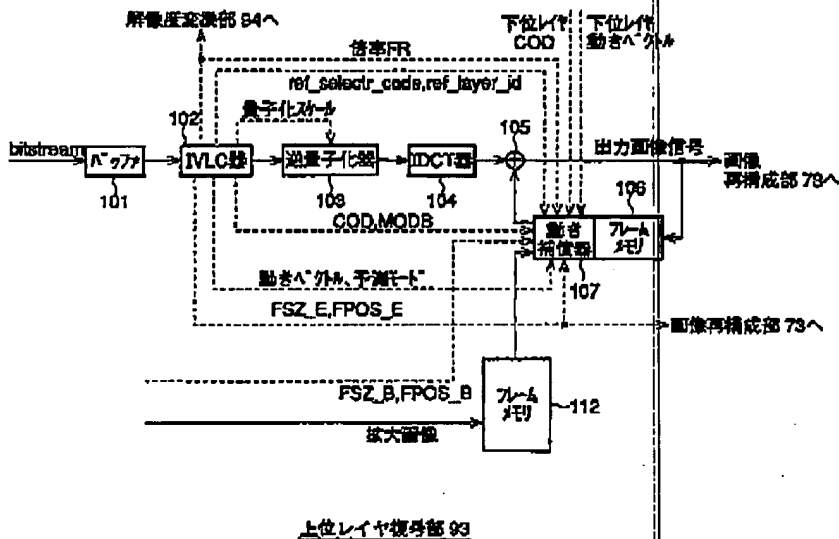
(40)

特開 2000-13790

【図 15】



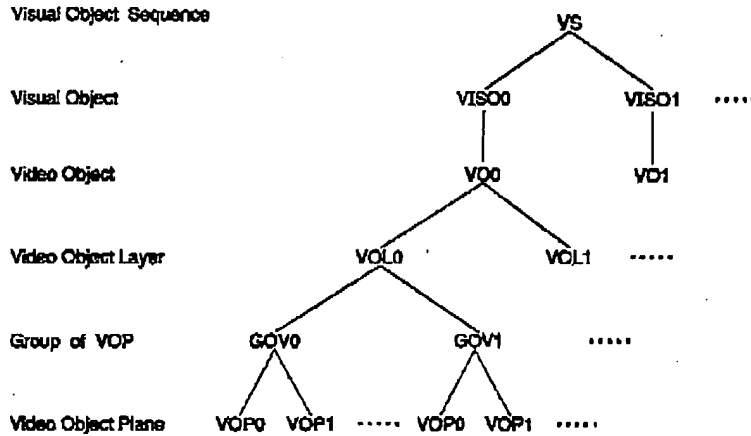
【図 16】



(41)

特開 2000-13790

【図 17】



【図 18】

Video Session Class		
Syntax	No. of bits	Mnemonic
VideoSession () {		
video_session_start_code	sc+8-32	
do {		
VideoObject ()		
while(nextbits_bytaligned 0—video_object_start_code)		
next_start_code ()		
video_session_end_code	sc+8-32	
}		

*concurrent loop solution to be provided by MSDL

sc

This code is intended to be used in combination with additional bits, for the purpose of the synchronization. Its binary representation is 23 zeros followed by a 1 (000000000000000000000001), or its hexadecimal representation is "000001".

video_session_start_code

This code cannot be emulated by any combination of other valid bits in the bitstream, and is used for synchronization purpose. Its value is that of the sc followed by "80" in hexadecimal.

video_session_end_code

This code cannot be emulated by any combination of other valid bits in the bitstream, and is used for synchronization purpose. Its value is that of the sc followed by "81" in hexadecimal. video_session_end_code results all data relative to VOPs. In other words, different session are issued completely independently.

【図 20】

Video Object Layer Class		
Video Object Layer		
Syntax	No. of bits	Mnemonic
VideoObjectLayer () {		
video_object_layer_start_code	sc+4-28	
video_object_layer_id	4	
video_object_layer_shape	2	
if(video_object_layer_shape == "00") {		
video_object_layer_width	13	
video_object_layer_height	13	
}		
not_8_bit	1	
obmc_disable	1	
if(not_8_bit) {		
quant_precision	4	
bits_per_pixel	4	
}		
video_object_layer_shape_effects	4	
if(video_object_layer_shape_effects == "0001")		
if(video_object_layer_shape_effects == "0011")		
if(video_object_layer_shape_effects == "0100")		
if(video_object_layer_shape_effects == "0101") {		
video_object_layer_feathering_disc	3	
if(video_object_layer_shape_effects == "0100")		
if(video_object_layer_shape_effects == "0101") {		
for(i=0; i<video_object_layer_feathering_disc; i++)		
feathering_filter ();	8*15	
}		
video_object_layer_sprite_usage		
if(video_object_layer_sprite_usage != SPRITE_NOT_USED) {	2	

【図 28】

VideoObjectPlane Class		
VideoObjectPlane		
Syntax	No. of bits	Mnemonic
VideoObjectPlane () {		
VOP_start_code	sc+8-32	
do {		
modulo_time_base	1	
while(modulo_time_base != "0")		
VOP_time_increment	1-15	
vop_codec	1	
if(vop_codec == "0") next_start_code ()		
VOP_prediction_type	2	
if(newpred_enable) {		
vop_id	4-15	
vop_id_for_prediction_indication	1	
if(vop_id_for_prediction_indication)		
vop_id_for_prediction	4-15	
marker_bit	1	
}		
intra_dc_vlc_for	3	
vop_reduced_resolution	1	
if(VOP_prediction_type == "P") {		
video_object_layer_sprite_usage != STATIC_SPRITE &&		
VOP_prediction_type == SPRITE)		
VOP_rounding_type	1	
if(video_object_layer_sprite_usage != SPRITE_NOT_USED)		
&&VOP_prediction_type == SPRITE) {		
if(nbr_of_sprites > 0) {		
encode_sprite_trajectory ()		
if(flighting_change_in_sprite) {		

(42)

特開 2000-13790

【図 21】

```

if(video_object_layer_sprite_usage==STATIC_SPRITE)
{
    video_object_layer_sprite_usage=ON_LINE_SPRITE;
    sprite_holm 13
    sprite_valm 13
}
if(video_object_layer_sprite_usage==STATIC_SPRITE)
{
    video_object_layer_sprite_usage=ON_LINE_SPRITE;
    sprite_Left_Edge 13
    sprite_Top_Edge 13
}
no_of_sprite_points 6
warping_accuracy 2
lighting_change_in_sprite 1
}
if(video_object_layer_shape== '00' )
{
    video_object_layer_shape_type 1
    video_object_layer_quant_type 1
}
if(video_object_layer_quant_type){
    load_intra_quant_mat 1
    if(load_intra_quant_mat){
        i=0;
        do
            intra_quant_mat[i] 6
            while(intra_quant_mat[i] !=0
                && i++<64);
            for(j=i;<64;j++)
                intra_quant_mat[j]=
                    intra_quant_mat[i-1];
        }
        load_nonintra_quant_mat 1
        if(load_nonintra_quant_mat){
            i=0;
            do
                nonintra_quant_mat[i] 6

```

【図 22】

```

while(nonintra_quant_mat[i] !=0
    && i++<64);
for(j=i;<64;j++)
    nonintra_quant_mat[j]=
        nonintra_quant_mat[i-1];
}
if(video_object_layer_shape== '10' ){
    disable_gray_quant_update 1
    load_gray_intra_quant_mat 1
    if(load_gray_intra_quant_mat){
        i=0;
        do
            gray_intra_quant_mat[i] 6
            while(gray_intra_quant_mat[i] !=0
                && i++<64);
            for(j=i;<64;j++)
                gray_intra_quant_mat[j]=
                    gray_intra_quant_mat[i-1];
        }
        load_gray_nonintra_quant_mat 1
        if(load_gray_nonintra_quant_mat){
            i=0;
            do
                gray_nonintra_quant_mat[i] 6
                while
                    gray_nonintra_quant_mat[i] !=0
                    && i++<64);
            for(j=i;<64;j++)
                gray_nonintra_quant_mat[j]=
                    gray_nonintra_quant_mat[i-1];
        }
    }
    quant_sample 1
    Complexity_estimation_disable 1
    if(Complexity_estimation_disable){

```

【図 23】

```

Parameter_list 6
Estimation_method 2
if(Estimation_method== '00' ){
    INTRA 1
    INTRA+Q 1
    INTER 1
    INTER+Q 1
    INTERPOLATE MC+Q 1
    FORWARD MC+Q 1
    BACKWARD MC+Q 1
    H.263 PB DIRECT 1
    Not_Coded 1
    Zg-zag_DCT_coeff 1
    Half_pel_advanced_prediction 1
    Half_pel_normal_pred 1
    VLC_symbol 1
    Scalability_parameters {
        t.b.d..... 1
    }
    Scalability_parameters {
        t.b.d..... 1
    }
    Sprite_coding_parameters {
        t.b.d..... 1
    }
}
if(Estimation_method== 'xx' ){
    t.b.d.....
}
}
error_resilent_disable 1
if(error_resilent_disable){
    data_partitioning 1
    reversible_VLC 1
    newpred_enable 1

```

【図 24】

```

if(newpred_enable)
    newpred_mode_flag 2
}
intra_code_pred_disable 1
separable_motion_shape_texture 1
if(video_object_layer_sprite_usage==STATIC_SPRITE){
    encodeSpritePlane 0
}
Scalability 1
if(scalability){
    ref_layer_id 4
    ref_layer_sampling_direct 1
    hor_sampling_factor_n 5
    hor_sampling_factor_m 5
    vert_sampling_factor_n 5
    vert_sampling_factor_m 5
    enhancement_type 1
}
do {
    VideoObjectPlane 0
    while(next_plane_byte_assigned !=video_object_plane_start_code)
        next_start_code 0
}

```

video_object_layer_start_code

This is a Unique code of length 28 bits (6+4) that precedes the video_object_layer_id.

video_object_layer_id

This is a 4-bit code which identifies a video object layer for a video object being processed.

video_object_layer_shape

This is a 2-bit code which identifies the shape type of video object layer as shown in Table 16.

video_object_layer_shape	Code
rectangular	00
binary	01
gray-scale	10
Binary-shape-only	11

Table 16 Video Object Layer shape types

(43)

特開 2000-13790

【図 25】

This flag is "00" if a VOL shape is rectangular, "01" if a VOL has a binary shape (i.e. if each pixel of the rectangle defined by video_object_layer_width and video_object_layer_height is either part of a VOL or not) and "10" if a VOL shape is defined by gray scale data (i.e. if each pixel of the rectangle defined by video_object_layer_width and video_object_layer_height is to be uniquely combined with the pixel of other VOLs at the same spatial location). The flag is "11" if the VOL consists of only binary shape and texture.

video_object_layer_width video_object_layer_height

These two codes define the picture size for the session, in pixels unit (zero values are forbidden).

This is also the size of the unique VOL of the session.

not_8bit

This one bit flag is set when the video data precision is not 8 bits per pixel.

quant_precision

This field specifies the number of bits used to represent quantizer parameters. Values between 3 and 9 are allowed.

When not_8bit is zero, and therefore quant_precision is not transmitted, it takes a default value of 6.

bits_per_pixel

This field specifies the video data precision in bits per pixel. It may take different values for different video object layer within a single video object. A value of 10 in this field would indicate 12 bits per pixel. This field may take values between 4 and 12.

When not_8bit is zero and bits_per_pixel is not present, the video data precision is always 8 bits per pixel, which is equivalent to specifying a value of 8 in this field.

video_object_layer_sprite_usage

This 2-bit code is to indicate the usage of sprite coding or global motion compensation (GMC).

0000 means that no sprite is used in the coding. Mnemonic is SPRITE_VOP_DISABLE.

0001 means that a sprite is being used directly (i.e., no prediction are made on the basis of the sprites covering some or all of the VOPs in this VOL (i.e., static sprite). The static sprite is assumed to have been transmitted as an I-VOP with possible additional pieces localized in the future also as I-VOPs and residual update pieces also transmitted as I-VOPs but using the quantization of non_intra macroblocks. Mnemonic is STATIC_SPRITE.

0010 means that a sprite is used in predictive coding for some or all of the VOPs in this VOL. The sprite is being constructed on-line and the mechanism for building the sprite is on. Mnemonic is ON-LINE_SPRITE.

0011 means that GMC is used in predictive coding for some or all of VOPs in this VOL. GMC is an extension of on-line sprite coding where blending factor is always 1 and the size of the sprite is just the same as that of the reference VOP. This means that the previous reconstructed VOP is regarded as an on-line sprite, and no additional memory for a sprite is needed. Mnemonic is GMC.

sprite_horiz

This 10-bit field defines the horizontal dimension of the sprite.

sprite_vert

This 10-bit field defines the vertical dimension of the sprite.

sprite_Left_Edge

This 10-bit field defines the left edge of the sprite.

sprite_Top_Edge

This 10-bit field defines the top edge of the sprite.

no_of_sprite_points

【図 27】

Time code	range of value	No. Of bits	Mnemonic
time code hours	0-23	5	uint5bit
time code minutes	0-59	6	uint6bit
marker bit	1	1	bit1
time code seconds	0-59	6	uint6bit

Table 19 time code

closed_gov

This is a one-bit flag which indicates the nature of the predictions used in the first consecutive B-VOPs (if any) immediately following the first coded I-VOPs after the GOV header. The closed_gov is set to '1' to indicate that these B-VOPs have been encoded using only backward prediction or intra coding. This bit is provided for use during any editing which occurs after encoding. If the previous pictures have been removed by editing, broken_link may be set to '1' so that a decoder may avoid displaying these B-VOPs following the first I-VOP following the group of plane header. However if the closed_gov bit is set to '1', then the editor may choose not to set the broken_link bit as these B-VOPs can be correctly decoded.

broken_link

This is a one-bit flag which shall be set to '0' during encoding. It is set to '1' to indicate that the first consecutive B-VOPs (if any) immediately following the first coded I-VOP following the group of plane header may not be correctly decoded because the reference frame which is used for prediction is not available (because of the action of editing). A decoder may use this flag to avoid displaying frames that cannot be correctly decoded.

【図 26】

Group Of VOPs class

Syntax of Group of VideoObjectPlane:

group of VideoObjectPlane {	No. Of bits	Mnemonic
group_start_code	32	uint32bit
VOP_time_increment_resolution	15	uint16bit
marker_bit	1	bit1
time_code	18	uint18bit
closed_gov	1	uint1bit
broken_link	1	uint1bit
do {		
group of VideoObjectPlane {		
while(nextbits_bytaligned()) {--group_start_code}		
next_start_code 0		

group_start_code

The group_start_code is the unique of length of 32bit. It identifies the beginning of a GOV header.

VOP_time_increment_resolution

The VOP_time_increment_resolution is a 15bit unsigned integer that indicates the resolution of the VOP_time_increment in terms of number of ticks within one modulo time (one second in this case). The value of zero is forbidden.

Time_code

This is a 18-bit integer containing the following: time_code_hours, time_code_minutes, marker_bit and time_code_seconds as shown in Table 19. The parameters correspond to those defined in the IEC standard publication 481 for "Time and control codes for video tape recorders". The time code refers to the first plane (i.e. display order) after the GOV header.

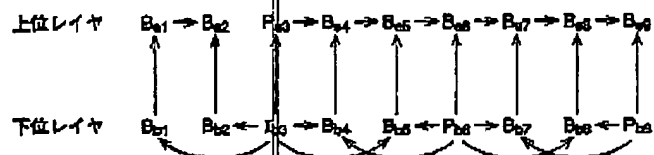
【図 29】

```

lighting_change_factor_encode 0
{
  if(video_object_layer_sprite_usage==STATIC_SPRITE){
    if(sprite_transmit_mode!=STOP){
      do{
        sprite_transmit_mode=checkSpriteTransMode 0
        sprite_transmit_mode
        if(sprite_transmit_mode==PIECE){
          (sprite_transmit_mode==UPDATE)
          encodeSpritePiece 0
        }while(sprite_transmit_mode!=STOP)
        if(sprite_transmit_mode!=PAUSE)
      }
    }else if(video_object_layer_sprite_usage==ON-LINE_SPRITE){
      blending_factor
    }
  }
  if(video_object_layer_shape != "00"){
    VOP_width
    marker_bit
    VOP_height
    VOP_horizontal_mc_spatial_ref
    marker_bit
    VOP_vertical_mc_spatial_ref
    if(video_object_layer_sprite_usage==STATIC_SPRITE){
      return 0
    }
    if(scalability && enhancement_type)
      background_composition
    VOP_CR
    change_CR_disable
  }
  if(interlaced)
    if(interlaced)
      top_field_first
    if(VOP_prediction_type=="10" && scalability)
  }
}

```

【図 45】



【图 30】

```

VOP_quant      2
else {
  VOP_quant      quant_precision
  R(video_object_layer_shape-- "10" )      6
    VOP_gray_quant
  }
  if (video_object_layer_shape_effects-- "0010" ) {
    (video_object_layer_shape_effects-- "0011" ) {
      (video_object_layer_shape_effects-- "0101" ) {
        VOP_constant_alpha      1
        R(VOP_constant_alpha)
        VOP_constant_alpha_value      3
      }
    }
    video_object_plane_fcode_forward      3
    video_object_plane_fcode_backward      3
    if (Complexity_estimation_disable) {
      if (estimation_method-- "00" ) {
        if (VOP_prediction_type-- "00" ) {
          in(INTRA)
            DCECS_intra      B
          if (INTRA+Q)
            DCECS_intra_Q      3
          if (Not_coded)
            DCECS_Not_coded      3
          if (zig-zag_DCT_coeff)
            DCECS_zig-zag_DCT      VLC
          if (half_pel_advanced_prediction)
            DCECS_half_pel_ap      3
          if (half_pel_normal_pred)
            DCECS_half_pel_norm_pred      3
          if (VLC_symbol)
            DCECS_VLC_symbol      3
          if (shape_coding_parameters) {
            DCECS_shape_coding      7bd
          }
        }
      }
    }
  }
}

```

【图 3-1】

if(VOP_prediction_type == '0') {	
if(estimation_method == '00') {	
in(INTRA)	B
DCECS_intra	B
in(INTRA+Q)	B
DCECS_intra_Q	B
if(Not coded)	B
DCECS_Not_coded	B
if(INTER)	B
DCECS_inter	B
if(INTER4V)	B
DCECS_inter4V	B
if(INTRA+Q)	B
DCECS_intra_Q	B
if(2g--zag_DCT_coeff)	VLC
DCECS_2g--zag_DCT	
if(Half_pel_advanced_prediction)	B
DCECS_half_pel_ap	B
if(Half_pel_normal_pred)	B
DCECS_half_pel_np	B
if(VLC_symbol)	B
DCECS_VLC_symbol	B
if(Shape_coding_parameters) {	
DCECS_LB.....	Tbd
}	
if(estimation_method == '00') {	
if(VOP_prediction_type == '0') {	
in(INTRA)	B
DCECS_intra	B
in(INTRA+Q)	B
DCECS_intra_Q	B
if(Not coded)	B
DCECS_Not_coded	B

【图 3 2】

#(INTERP		
DCEGS_inter		8
#(INTER4V)		
DCEGS_inter4V		8
#(INTRFA+Q)		
DCEGS_intra_Q		8
#(INTERPOLATE MC+Q)		
DCEGS_interpolate		8
#(FORWARD MC+Q)		8
DCEGS_forward		8
#(BACKWARD MC+Q)		8
DCEGS_backward		8
#(H.263 PB DIRECT)		VLD
DCEGS_direct_h263		8
#(Zig-zag_DCT_coeff)		8
DCEGS_zig-zag_DCT		8
#(half_pel_advanced_prediction)		
DCEGS_half_pel_ap		8
#(half_pel_normal_pred)		
DCEGS_half_pel_np		Tbd
#(VLC_symbol)		
DCEGS_vlc_symbol		
#(shape_coding_parameters){		
DCEGS_..._t.b.d....		
}		
}		8
#(estimation_method-- "ad") {		
#(MCP_prediction_type-- "1") {		8
#(INTRFA)		
DCEGS_intra		8
#(INTRFA+Q)		
DCEGS_intra_Q		
#(Not coded)		
DCEGS_Not_coded		

【图 3-3】

```

(KINTER)
  DCECS_inter 8
W(INTERV)
  DCECS_interw 8
W(INTER+Q)
  DCECS_inter_Q 8
W(INTERPOLATE MO+Q)
  DCECS_interpolate 8
W(FORWARD MO+Q)
  DCECS_forward 8
W(BACKWARD MO-Q)
  DCECS_backward 8
W(H.263 PB DIRECT)
  DCECS_direct_h263 VLC
W(Zig-zag_DCT_coeff)
  DCECS_zigzag_DCT 8
R(half_pel_advanced_prediction)
  DCECS_half_pel_adv 8
R(half_pel_normal_pred)
  DCECS_half_pel_no 1b.d
I(VLC_symbol)
  DCECS_VLC_symbol 1b.d
W(shape_coding_parameters) {
  DCECS_shape...1b.d.....
}
W(sprite_coding_parameters) {
  DCECS_sprite...1b.d.....
}
}
}
W(availability) {
  W(temporal_motion_shape_texture)
  W(error_resilience_flag)
  combined_motion_shape_texture_coding()
  misc
}

```

(45)

特開2000-13790

【図34】

```

do {
  do {
    combined_motion_shape_texture_coding ()
    while(nextbits_bytaligned () != 0000 0000 0000 0000) {
      next_resync_marker ()
      resync_marker 17
      macroblock_number 1-12
      quant_scale quant_precision
      header_extension_code 1
      if(header_extension_code != 1) {
        do {
          module_time_base 1
          while(module_time_base != 0)
            VOP_time_increment 1-15
        }
      }
    }
    while(nextbits_bytaligned () != 000 0000 0000 0000 0000 0000)
  } else {
    if(video_object_layer_shape != "00") {
      do {
        first_shape_code 1-3
        while(count of macroblocks != total number of macroblocks)
      }
      if(error_resilience_disable) {
        motion_coding ()
        if(video_object_layer_shape != "00")
          shape_coding ()
        texture_coding ()
      }
    }
    do {
      do {
        motion_coding ()
        while(next_bits () != "1010 0000 0000 0000 1")
        motion_marker 17
      }
    }
  }
}

```

【図36】

```

marker_bit 1
forward_shape_vertical_mc_spatial_ref 13
forward_shape_coding ()
}
}
ref_select_code
if(VOP_prediction_type == "01" || VOP_prediction_type == "10") {
  forward_temporal_ref 1-15
  if(VOP_prediction_type == "10") {
    marker_bit 1
    backward_temporal_ref 1-15
  }
}
combined_motion_shape_texture_coding ()
}
do {
  VideoObjectPlane ()
  while(nextbits_bytaligned () != video_object_plane_start_code)
    next_start_code ()
}

```

VOP_start_code
This code cannot be associated by any combination of other valid bits in the bitstream, and is used for synchronization purposes. It is so followed by a unique 8-bit code.

module_time_base
This value represents the local time base at the one second resolution unit (1000 milliseconds). It is represented as a marker transmitted in the VOP header. The number of consecutive "1" followed by a "0" indicates the number of seconds has elapsed since the synchronization point marked by the module_time_base of the last displayed IP-VOPs belonging to the same VCL. There are two exceptions, one for the first IP-VOPs after the GOV header, and the other is for B-VOPs prior to display order to the first I-VOP after the GOV header.

For the first IP-VOP after the GOV header, the module_time_base indicates the time_code in the GOV header.

For the first B-VOPs prior to display order to the first I-VOP after the GOV header, the module_time_base indicates the time relative to the time_code in the GOV header.

Note: When the bitstream contains B-VOPs, the decoder needs to store two time_base, one is indicated by the last displayed IP-VOP or the GOV header and the other is indicated by the last decoded IP-VOP.

VOP_time_increment
This value represents the absolute VOP_time_increment from the synchronization point marked by the module_time_base measured in the number of clock ticks. It can take a value in the range [0, VOP_time_increment_resolution). The number of bits representing this value is calculated as the minimum number of bits required to represent the above range. The local time base in units of seconds is recovered by dividing this value by the VOP_time_increment_resolution. For IP and B-VOP this value is the absolute VOP_time_increment from the synchronization point marked by the module_time_base.

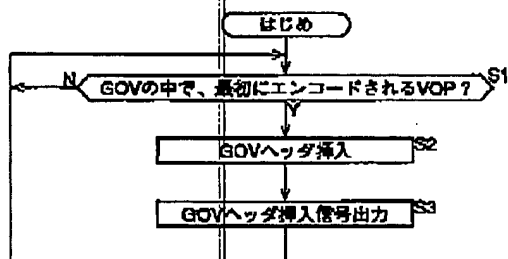
【図35】

```

if(video_object_layer_shape != "00")
  shape_coding ()
do {
  texture_coding ()
  while(nextbits_bytaligned () != "0000 0000 0000 0000") {
    if(nextbits_bytaligned () != 000 0000 0000 0000 0000 0000) {
      next_resync_marker ()
      resync_marker 17
      macroblock_number 1-12
      quant_scale 5
      header_extension_code 1
      if(header_extension_code != 1) {
        do {
          module_time_base 1
          while(module_time_base != 0)
            VOP_time_increment 1-15
        }
      }
    }
    while(nextbits_bytaligned () != 000 0000 0000 0000 0000 0000)
  }
} else {
  if(enhancement_type) {
    load_backward_shape 1
    if(load_backward_shape) {
      backward_shape_width 13
      backward_shape_height 13
      backward_shape_horizontal_mc_spatial_ref 13
      marker_bit 1
      backward_shape_vertical_mc_spatial_ref 13
      backward_shape_coding ()
    }
    load_forward_shape 1
    if(load_forward_shape) {
      forward_shape_width 13
      forward_shape_height 13
      forward_shape_horizontal_mc_spatial_ref 13
    }
  }
}

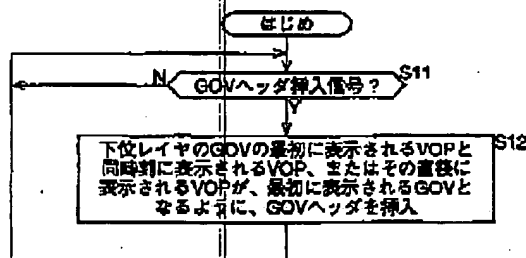
```

【図43】



下位レイヤのVLC処理

【図44】

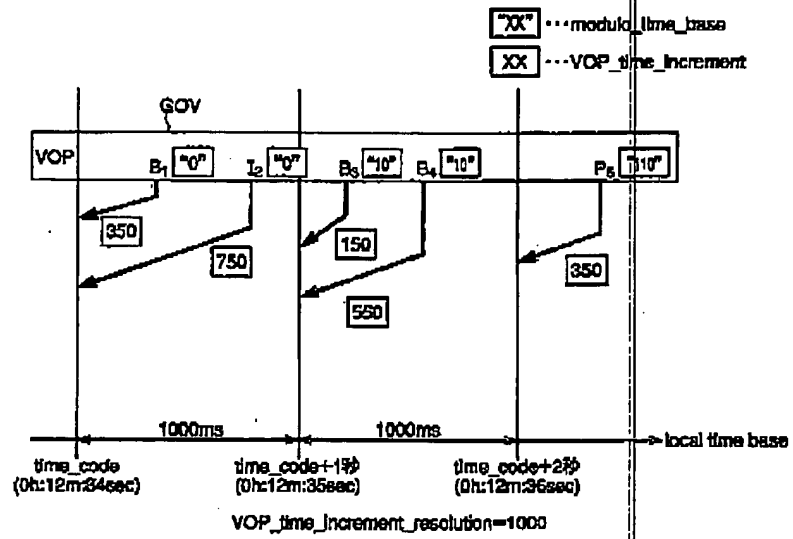


上位レイヤのVLC処理

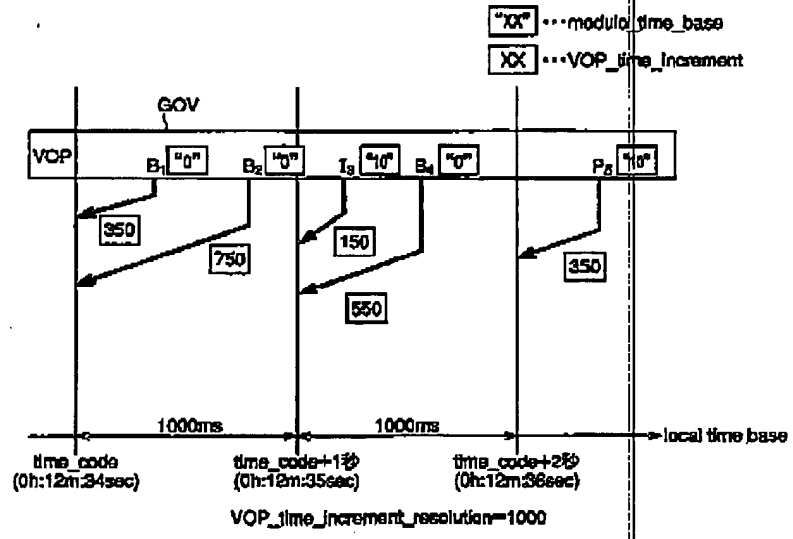
特開 2000-13790

(46)

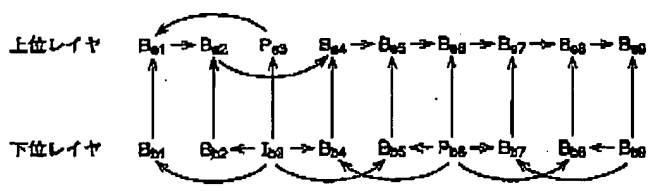
【図 37】



【図 38】



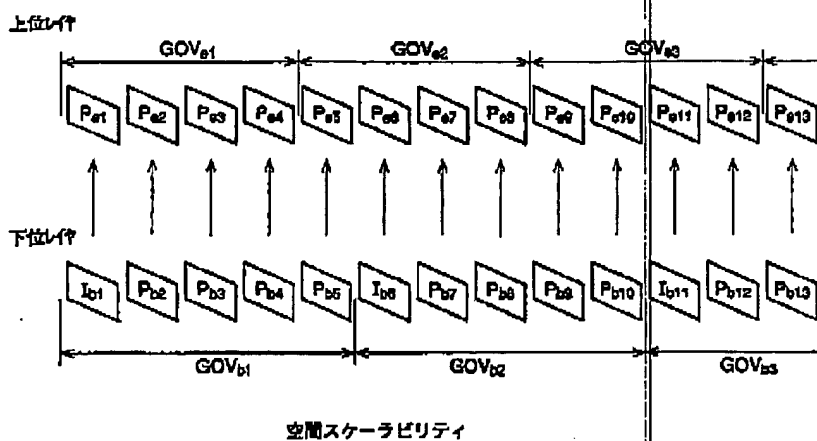
【図 49】



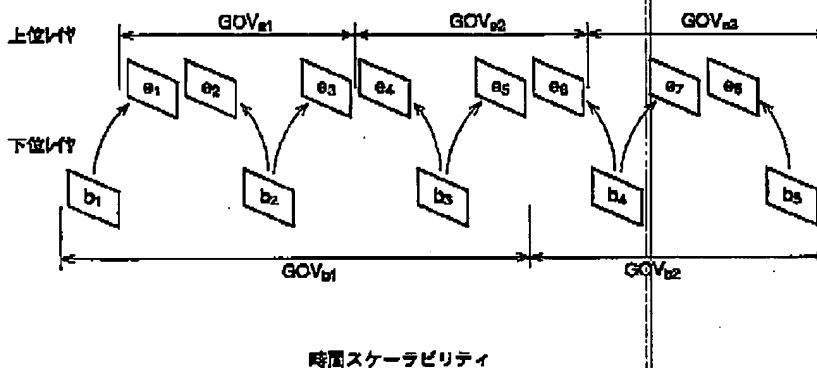
特開 2000-13790

(47)

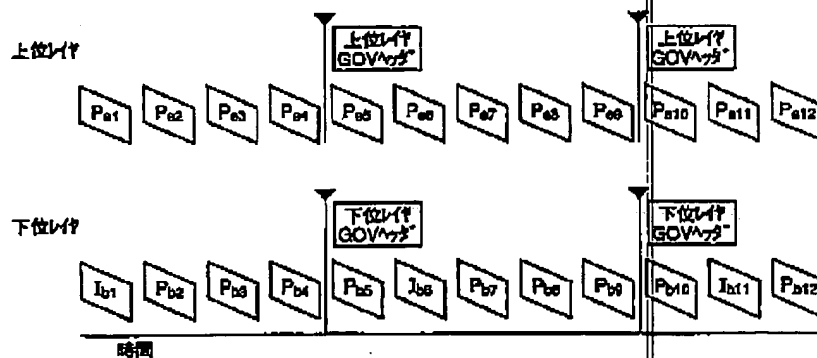
【図 39】



【図 40】



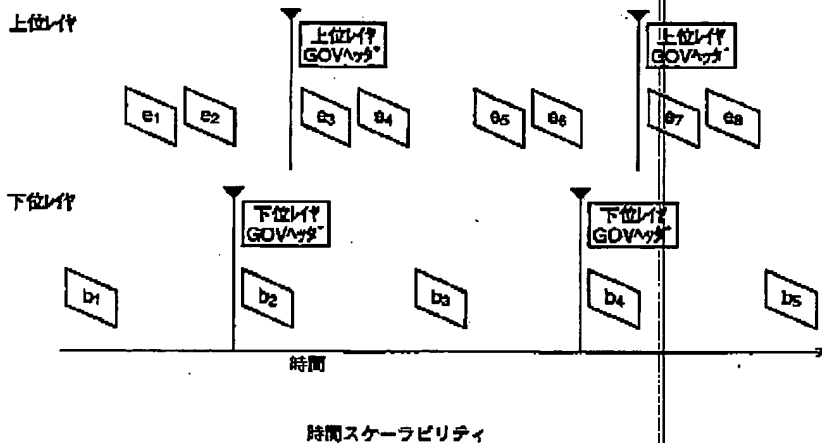
【図 41】



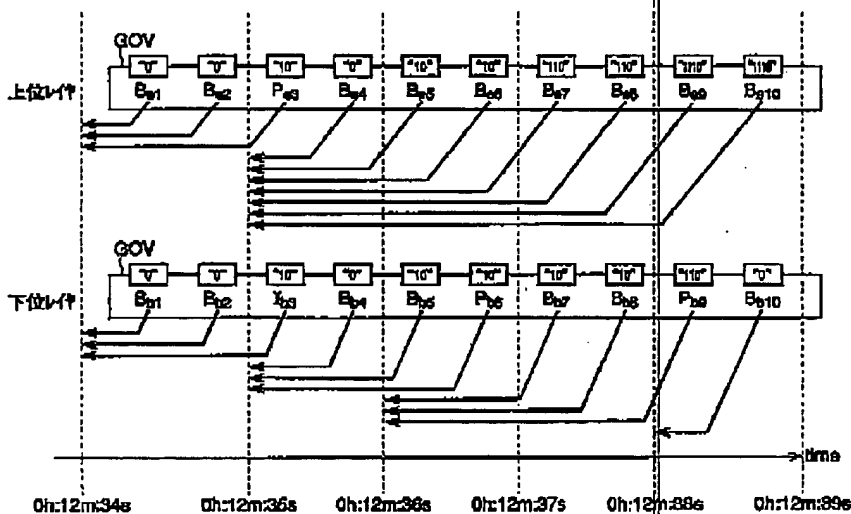
特開2000-13790

(48)

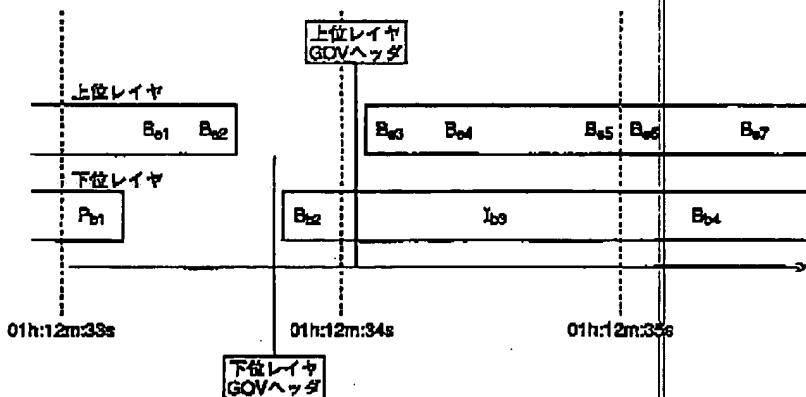
【図42】



【図46】



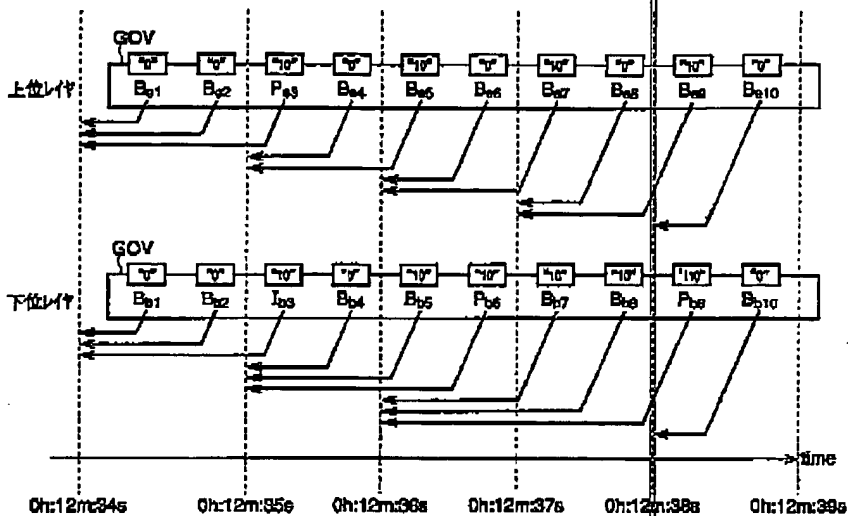
【図50】



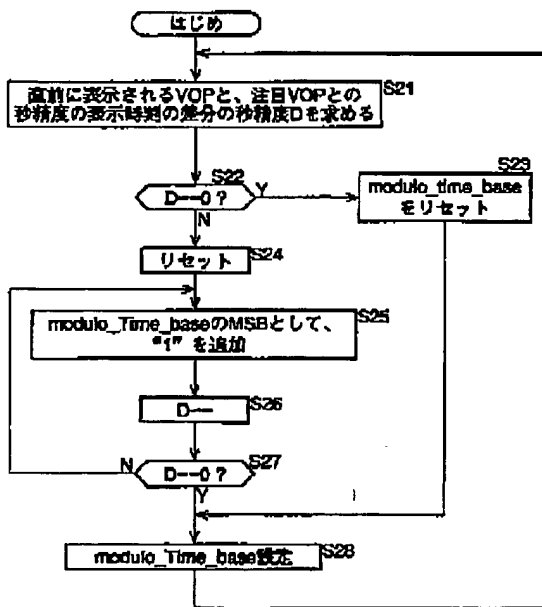
(49)

特開2000-13790

【図47】



【図48】

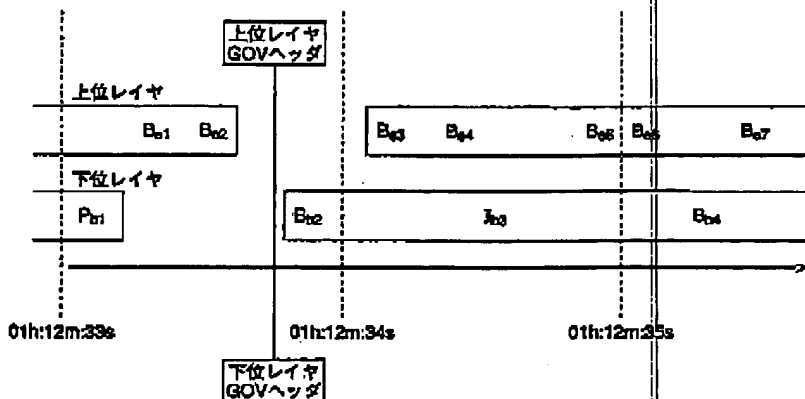


上位レイヤのVOPについてのmodule_time_baseの設定

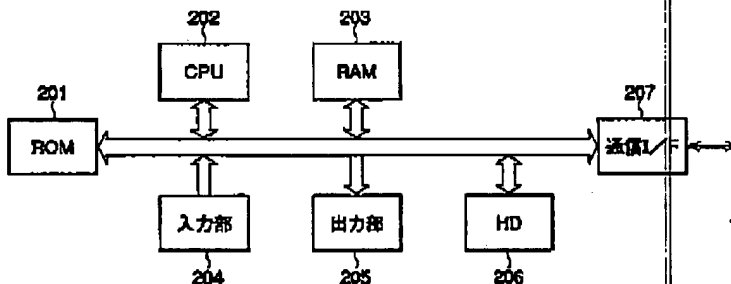
(50)

特開 2000-13790

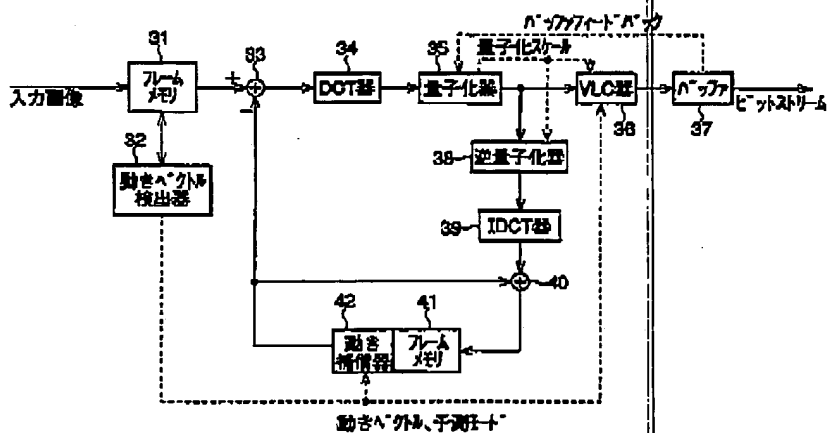
【図 5 1】



【図 5 2】



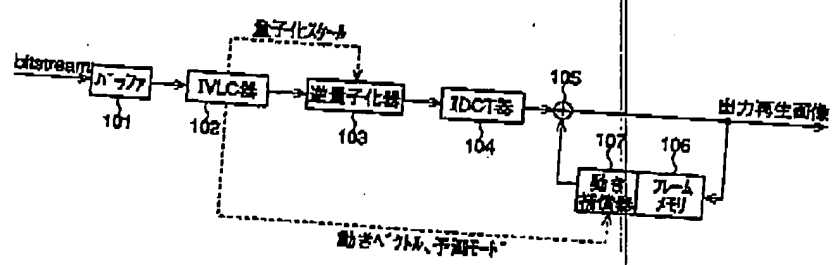
【図 5 3】



特開2000-13790

(51)

【図54】



フロントページの続き

(72)発明者 矢ヶ崎 陽一
東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

Fターム(参考)

5C059 KK00 MA00 MA23 MA31 MB12
MB22 NN01 PP04 PP05 PP06
PP07 RB01 RB18 RC04 RC19
SS07 TA11 TB03 TC36 TD05
UA02 UA03 UA05 UA06
5C078 BA57 BA64 DA00 DA01 DA02
EA07

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.